А.И. МИХАЙЛОВ
А.И. ЧЕРНЫЙ
Р.С. ГИЛЯРЕВСКИЙ

# ОСНОВЫ ИНФОРМАТИКИ

*А. И. Михайлов, А. И. Черный,*
*Р. С. Гиляревский*

6.02
М 69

И. 1992.

# ОСНОВЫ ИНФОРМАТИКИ

*Второе, переработанное и дополненное издание*

174028

БИБЛИОТЕКА
Сан. СХИ
гор. Самарканд

323

ИЗДАТЕЛЬСТВО «НАУКА»
*Москва 1968*

*Информатика — новая научная дисциплина, изучающая структуру и свойства научной информации, закономерности научно-информационной деятельности, ее теорию, историю, методику и организацию. В книге формулируется предмет и метод информатики, дается характеристика различных видов документов как источников научной информации, освещаются методы аналитико-синтетической переработки документов, излагаются принципы информационного поиска и пути его механизации и автоматизации, описываются средства оперативного копирования и размножения документов. Большое внимание уделяется раскрытию принципов организации научно-информационной деятельности в СССР и за рубежом.*

*Первое издание книги (1965 г.) под названием «Основы научной информации» вызвало много откликов в отечественной и зарубежной печати. В настоящее издание введены новые фактические данные, исправлены некоторые неточности. Отдельные главы подверглись коренной переработке.*

6-10
677—68 (I)

# ПРЕДИСЛОВИЕ
# К ПЕРВОМУ ИЗДАНИЮ

Перед читателем — книга, созданная Всесоюзным институтом научной и технической информации. Среди сотен тысяч страниц — плодов деятельности института — это первое детище, посвященное предмету его деятельности, научной информации, подведению итогов мирового и собственного опыта.

Вряд ли будет ошибкой сказать, что в области научной информации ВИНИТИ концентрирует бóльшую часть опыта нашей страны. Для всех интересующихся этим вопросом, — а какой же работник науки им не интересуется, — выход в свет этой книги не безразличен.

Институт научной информации был организован по представлению Академии наук СССР в 1952 г., а позднее перешел в совместное подчинение Академии наук СССР и Государственного комитета Совета Министров СССР по новой технике (Гостехники) и стал называться Всесоюзным институтом научной и технической информации, значительно расширив свои функции. Ныне роль Гостехники в отношении ВИНИТИ взял на себя Государственный комитет Совета Министров СССР по науке и технике.

ВИНИТИ представляет собой теперь громадную «обогатительную фабрику» научной литературы. Одновременно он вынужден существенную часть своих усилий тратить на разработку самих методов научной информации и изучение того, как поставлено это дело в мире. Здесь не только открывается огромный простор, но и возникает абсолютная необходимость возможной (малой и большой) механизации и автоматизации всего «обогатительного процесса», нашедшего пока в деятельности ВИНИТИ выражение в виде издания ряда реферативных журналов, стремящихся исчерпать мировую периодику в области естественных наук и техники, а также в виде выборочной «экспресс-информации» и обобщающих «итогов науки».

Насколько это нужное дело? Все в большей и быстро возрастающей степени, но в разной для разных дисциплин. Математик имеет дело с гораздо более ограниченным кругом периодики, и статьи его благодаря лаконичности математической символики по сути уже являются рефератами. Это ограничивает его интерес к рефе-

ративному журналу. Другим полюсом является химик: интересующие его сведения разбросаны по множеству специальных химических и еще большему множеству нехимических журналов.

По ориентировочным данным, в настоящее время в мире выходит около 10 тыс. названий периодических изданий, представляющих непосредственный интерес для химиков. В этих периодических изданиях ежегодно помещается не менее 200 тыс. публикаций по химии и химической технологии. Кроме того, в мире публикуется до 5 тыс. книг, свыше 30 тыс. патентов и около 20 тыс. научно-технических отчетов, полностью или частично относящихся к химии и химической технологии. В 1964 г. в реферативных журналах «Химия» и «Биологическая химия», выпускаемых ВИНИТИ, было освещено (в виде рефератов, аннотаций и библиографических описаний) свыше 150 тыс. статей, патентов, книг, научно-технических отчетов и других публикаций по профилю этих журналов.

Если бы химик, свободно владеющий 30 языками (условие невероятное), начал с 1 января 1964 г. читать все выходящие в этом году публикации, представляющие для него профессиональный интерес, и читал бы их по 40 часов в неделю со скоростью 4 публикации в час, то к 31 декабря 1964 г. он прочитал бы лишь $^1/_{20}$ часть этих публикаций.

При этом не следует забывать, что количество литературы по химии и химической технологии быстро растет, так как сейчас около 40% всех научных исследований и опытно-конструкторских работ ведется именно в этой области. За последние 15 лет годовой прирост химической литературы составлял несколько более 8,5%. При таком темпе роста количество химической литературы должно удваиваться примерно каждые 10 лет. Опыт работы ВИНИТИ и реферативного журнала «Chemical Abstracts», выходящего в США с 1907 г., полностью подтверждает эти оценки. К этому нужно добавить, что ежегодно в советских реферативных журналах «Химия» и «Биологическая химия» публикуется информация более чем о 180 тыс. химических соединений, из которых не менее 90 тыс. предположительно являются новыми. Всего к настоящему времени собрана информация примерно о 2,5 млн. органических и неорганических соединений.

Химик уже не в состоянии работать без реферативных журналов и их указателей (предметного, формульного, патентного, авторского), ежегодных и пятилетних, которые позволяют ему совершать экскурсы в литературу на десятки лет назад и вместе с тем быть в курсе литературы сегодняшнего дня. Вот почему реферативные журналы впервые появились в области химии, и химики пользуются ими более ста лет.

Количество литературы быстро растет и в других областях науки и техники. Поэтому современный исследователь все больше нуждается в помощи специальной службы, призванной ориенти-

ровать его в литературе. Такой службой и являются современные органы научной информации. Конечно, научно-информационная служба никогда не освободит исследователя от необходимости читать научно-техническую литературу в оригинале, но она может и должна избавить исследователя от траты времени на поиск и сбор информации. Дело информационных органов — поиск и концентрация (функция геолога и инженера-обогатителя), а усвоение — «выплавка металла» — всегда дело самого ученого.

Быстрое развитие научно-информационной деятельности с неизбежностью ведет к ее обособлению и выделению в самостоятельную область. Научно-информационная деятельность существует так же давно, как и сама наука. Но до последнего времени ее методы и средства разрабатывались в сфере разных научных дисциплин. Не случайно поэтому, что при формулировании своего предмета и метода научная информация широко использует достижения других наук.

При чтении данной книги бросается в глаза, что научная информация захватывает три совершенно различные области человеческой деятельности. Первая — это малоисследованная область творческого мышления человека и организации умственного труда. Проблема поиска необходимых специалисту сведений может быть успешно решена лишь при условии, что будут изучены логические операции, производимые ученым при поиске нужной ему информации. Она связана с такими разными вопросами, как, например, классификация наук, создание специальных языков науки, научный перевод и др. Примером может служить проблема оптимальной системы указателей. Теория научной информации, очевидно, смыкается с проблемой рациональной организации научного труда.

Вторая часть предмета научной информации — это довольно широкий комплекс вопросов, связанных с различными научными документами. Существующие виды научных документов сложились эмпирически в результате векового опыта. До сих пор не исследовано, насколько они отвечают требованиям современной науки. Некоторые сведения, приводимые в книге, позволяют думать, что ученые пока не располагают иными возможностями для распространения и сохранения во времени результатов своего труда. Все предложения относительно замены существующей системы научных публикаций являются в значительной мере умозрительными. Вопрос, однако, становится все более острым.

Третий комплекс вопросов, являющихся объектом исследования теории научной информации, относится к созданию технических средств, необходимых для осуществления успешной научно-информационной деятельности. Известно, что подготовка ряда зарубежных информационных изданий, например «Chemical Titles», осуществляется автоматически при помощи электронных машин. На этот же путь становится ВИНИТИ, начиная составлять

указатели к реферативным журналам при помощи новейших технических средств. Вряд ли приходится сомневаться, что в будущем огромная роль должна принадлежать электронным информационным машинам, обладающим колоссальным объемом долговременной памяти и способным быстро выдавать по запросам копии хранимой в их памяти литературы.

Все это требует достаточно длительной разработки, и данная книга скорее ставит вопросы, чем решает их. Несмотря на неизбежные недостатки этого первого опыта, полагаю, что книга будет полезна всем интересующимся вопросами научной информации.

*Академик А. Н. Несмеянов*

*9 августа 1965 г.*

# ПРЕДИСЛОВИЕ АВТОРОВ КО ВТОРОМУ ИЗДАНИЮ

Первое издание этой книги под заглавием «Основы научной информации» вышло осенью 1965 г. Оно явилось одной из первых попыток изложения на русском языке теории, методики и организации новой научной дисциплины.

Многочисленные отклики читателей подтвердили мнение, высказанное авторами в первом издании, что существует настоятельная необходимость осмыслить и обобщить факты, накопленные в результате многолетней научно-информационной деятельности, определить предмет и методы новой дисциплины, ее границы, место в ряду других наук, а также основные направления ее дальнейшего развития. К сожалению, до сих пор наша попытка создать обобщающий труд по этим вопросам остается единственной в отечественной литературе. Незначительный тираж первого издания (6800 экз.) и растущая потребность в пособиях, учебниках, руководствах, монографиях по этим вопросам вынуждают нас к переизданию данной книги, хотя мы хорошо сознаем ее неполноту и несовершенство.

В отзывах и рецензиях на первое издание были отмечены многие общие и частные недостатки по существу трактуемых нами проблем и по форме их изложения. Особенно серьезными представляются нам упреки, высказанные в рецензиях профессора Я. Г. Дорфмана (Институт истории естествознания и техники АН СССР) [1], сотрудников Государственной публичной научно-технической библиотеки [2], английского библиотековеда Г. Шура [3], а также в выступлениях при обсуждении книги на заседании лектория Государственной публичной библиотеки им. М. Е. Салтыкова-Щедрина в Ленинграде 7 февраля 1966 г. и на конференции в Новосибирском научном центре, организованной Государственной публичной научно-технической библиотекой Сибирского отделения АН СССР 15 июня 1966 г.

Одно из важных замечаний касалось названия книги и соответственно той научной дисциплины, которой книга посвящена.

---

[1] «Вестник АН СССР», 1966, № 11, стр. 102—103; «Научно-техн. информация», 1966, № 7, стр. 46—47.

[2] В. В. Власов, А. В. Кременецкая, В. Б. Петровский, Н. Г. Старобинская. Предмет исследования — научная информация. «Технические библиотеки СССР», 1966, вып. 3, стр. 44—50.

[3] «Journal of Documentation», 1966, v. 22, № 1, стр. 63—66.

Неудовлетворенность многих специалистов хорошо выразил проф. Я. Г. Дорфман, написавший, что это название крайне неудачно, так как речь идет о принципах и практических методах научно-информационной деятельности, а отнюдь не о самой научной информации. Внесенное им предложение назвать дисциплину «информатикой» не было первым подобным предложением, но стало последней каплей, перевесившей чашу весов в пользу этого предложения. Мы приняли его, и это отражено в заглавии и тексте второго издания книги. Есть основания надеяться, что новое название привьется не только у нас в стране, но и за рубежом и станет международным термином.

Другим серьезным недостатком первого издания рецензенты сочли неодинаковую степень обобщения материала. Действительно, в ряде глав и разделов отсутствовали сравнительный анализ и критика приводимых принципов, методов и средств информатики, изложение носило описательный характер, а в некоторых случаях — следы индивидуального стиля каждого автора. Мы постарались устранить этот важный недочет и сделать книгу более стройной по композиции и ровной по стилю. Однако мы отчетливо сознаем, что полностью избежать этого недостатка нам не удалось.

Многих рецензентов не удовлетворил объем вопросов, освещенных в первом издании: одним казалось, что он слишком широк, другим, что слишком узок. Работники библиотек, например, возражали против включения в книгу некоторых сведений из книговедения, библиотековедения и библиографии, которые, по их мнению, могут быть почерпнуты из учебников и монографий по данным дисциплинам. Практические работники служб научной информации, напротив, указывали на необходимость расширения этих и других разделов книги за счет более детального освещения методических и организационных вопросов научно-информационной и библиографической работы. Они предлагали, в частности, больше внимания уделить опыту организации БТИ и ОНТИ на предприятиях и в научно-исследовательских учреждениях, методике библиографического описания, классификации, реферирования, перевода научно-технической литературы, вопросам практической транскрипции и транслитерации и т. п.

Больше того, научные работники, аспиранты и студенты вузов, которые пользуются книгой как потребители информации, хотели бы видеть в ней краткое и доступное изложение теории множеств, необходимое для понимания проблем информационного поиска. Многие из них предложили включить во второе издание раздел по рациональной организации научного труда.

Эти часто противоречивые требования в известной мере вызваны несколько неопределенным читательским назначением первого издания, сформулированным следующим образом: «Книга рассчитана на широкие круги специалистов, работающих в области

научной информации и использующих ее в своей научной и практической работе. . . Ввиду отсутствия учебных пособий и практических руководств по научной информации монография может быть использована также студентами высших учебных заведений и лицами, готовящимися работать в области научной информации».

Авторы отдают себе полный отчет в том, что для названного выше широкого круга читателей нужна не одна-единственная книга, а целая система разных книг. Это — практические пособия для научных работников, инженеров и других специалистов-практиков, являющихся потребителями научной информации, учебные пособия для студентов вузов, т. е. будущих потребителей информации, руководства для специалистов и практических работников научно-информационных служб различного уровня и, наконец, обобщающие теоретические труды для специалистов, разрабатывающих отдельные проблемы информатики.

Собственно говоря, данная монография в первую очередь ориентирована именно на этот последний круг читателей. Однако, поскольку она неизбежно должна будет восполнять отсутствие других перечисленных типов книг, трудность, возникшая при выпуске первого издания, а именно слишком широкое читательское назначение, стоит перед нами и сейчас. Именно поэтому во втором издании книги сохранены и даже несколько расширены сведения из смежных дисциплин, необходимые для понимания проблем информатики.

В новом издании исправлены допущенные ранее неточности и восполнены некоторые пробелы, некоторые вопросы изложены по-новому. В первой главе уточнена терминология, относящаяся к основным понятиям информатики, и дается подробное обоснование этого нового термина. Переработана глава об информационно-поисковых системах (которая во втором издании разделена на три главы: «Информационно-поисковые системы», «Традиционные информационно-поисковые языки» и «Дескрипторные информационно-поисковые языки»), а также глава о репрографии.

Значительной переделке подверглась глава об организации информационной деятельности. Выход постановления Совета Министров СССР «О единой общегосударственной системе научно-технической информации» от 29 ноября 1966 г. четко определил функции различных органов информации в нашей стране и наметил дальнейшую программу развития научно-информационной деятельности в СССР. Это позволило по-новому изложить организационные проблемы, стоящие перед советскими службами научно-технической информации.

Авторы приносят глубокую благодарность всем, кто своими критическими замечаниями облегчил их труд по подготовке второго издания книги.

*Декабрь 1967 г.*

Глава 1

# ПРЕДМЕТ И МЕТОД ИНФОРМАТИКИ

## РОЛЬ НАУКИ В СОВРЕМЕННОМ ОБЩЕСТВЕ

Наука играет исключительно важную роль в строительстве коммунистического общества. Наука является одним из основных инструментов познания объективного мира. Наука играет важную роль в формировании мировоззрения людей. Без учета достижений науки невозможно крупное промышленное производство, невозможно планирование народного хозяйства. Именно поэтому Коммунистическая партия Советского Союза всегда уделяла огромное внимание развитию советской науки. В Программе КПСС, принятой XXII съездом нашей партии, говорится: «Партия будет всемерно содействовать дальнейшему **усилению роли науки** в строительстве коммунистического общества, поощрению исследований, открывающих новые возможности в развитии производительных сил, широкому и быстрому внедрению в практику новейших научно-технических достижений, решительному подъему экспериментальных работ, в том числе непосредственно на производстве. . . Наука станет в полной мере непосредственной производительной силой» [1]. И далее: «Применение науки становится решающим фактором могучего роста производительных сил общества. Развитие науки и внедрение ее достижений в народное хозяйство будет и в дальнейшем предметом особой заботы партии» [2].

Важную роль науки подчеркнул также XXIII съезд КПСС, определивший перспективы и направления очередного этапа строительства коммунизма. В резолюции съезда отмечается, что деятельность советских ученых должна быть направлена на дальнейшее решение актуальных научных проблем современности, на всемерное ускорение научно-технического прогресса, быстрейшее внедрение результатов научных исследований в народное хо-

[1] Программа КПСС. М., Госполитиздат, 1962, стр. 159—160.
[2] Там же, стр. 269.

зяйство, на обеспечение высоких темпов роста производительности труда.

А. Н. Косыгин в речи на Всесоюзном совещании научных работников в Кремле (июнь 1961 г.) подчеркивал, что в наше время наука не только является основой любого производственного процесса, но и сама научная деятельность приобрела в известном смысле производственный характер. «При всей сложности и многообразии взаимосвязей науки и техники с материальным производством очевидно, что для всестороннего развития материального производства темпы развития техники должны превосходить темпы роста производства, а наука должна развиваться быстрее, чем развивается техника. Это может быть достигнуто лишь в условиях социалистических производственных отношений, при которых научно-технический прогресс ведет к облегчению труда, к сокращению рабочего времени, росту благосостояния народа и служит укреплению дела мира» [3].

Развитие науки привело к революционным преобразованиям в технике: к широкой автоматизации производственных процессов, применению новейших физико-химических методов в технологии, электронных вычислительных машин в управлении производством и т. п. Научные исследования закладывают базу техники будущего.

Исследования в области ядерной физики уже привели к расширению нашей энергетической базы. Наряду с топливом органического происхождения для получения энергии мы используем ядерное топливо. Идет интенсивное строительство атомных электростанций и кораблей с атомными двигателями. Широким фронтом развернуты исследования в области термоядерного синтеза. Эти работы раскрывают еще бóльшую перспективу в обеспечении человечества энергией в будущем. Достигнуты заметные успехи в преобразовании различных видов энергии непосредственно в электрическую энергию.

Исследования в области физики твердого тела и оптики привели к открытию квантово-механических генераторов света — лазеров. Это открытие, которое еще совсем в недавнем прошлом было достоянием только физических лабораторий, в удивительно короткий срок стало основой совершенно новых и разнообразных технических средств. Мощный световой пучок используется в приборостроении для сверления тончайших отверстий в сверхтвердых телах, например в алмазе. Лазеры используются в медицине для проведения тончайших глазных операций. Интенсивно разрабатываются методы использования лазерного пучка для связи.

Нельзя не отметить еще одно из технических приложений физики твердого тела. Мы имеем в виду так называемую твердотель-

[3] А. Н. Косыгин. За тесную связь науки с жизнью. В кн.: «Всесоюзное совещание научных работников в Кремле 12—14 июня 1961 г. (стенограммы, материалы)». М., 1961, стр. 33.

ную электронику. Всем теперь хорошо известно, что миниатюрные транзисторы вытесняют старые вакуумные электронные лампы. Твердотельные выпрямители постепенно вытесняют громоздкие газоразрядные приборы. Твердые электронные приборы и установки все больше и больше применяются вместо вакуумных почти во всех областях измерительной техники.

Другая область физики — физика низких температур — также шагнула из научных лабораторий в технику. Характерное для физики низких температур явление сверхпроводимости теперь широко используется в самых различных областях техники. Сверхпроводящие магниты, потребляющие ничтожное количество энергии, применяются в настоящее время для самых различных целей. Здесь имеются и уникальные установки для получения исключительно сильных магнитных полей (до 500 000 *гаусс*) и компактные установки для различных физических экспериментов. Еще одним важным применением физики низких температур в технике является создание миниатюрных узлов вычислительных машин. Применение явления сверхпроводимости в технике возрастает с каждым днем.

На новой основе сейчас развивается учение о прочности материалов. Еще совсем недавно это была эмпирическая наука, оперировавшая неточными исходными данными. Теперь наука о сопротивлении материалов имеет твердую научную базу в виде теории дислокаций. Эта теория помогла экспериментаторам создать исключительно прочные нитевидные кристаллы, так называемые «усы», прочность которых приближается к максимальной теоретически возможной. Мы знаем теперь законы роста кристаллов и умеем направленно изменять этот рост. Мы можем теперь получать кристаллы и некристаллические вещества с ничтожным количеством примесей. Во многих случаях оказалось, что такие исключительно чистые материалы обладают уникальными свойствами.

Приведенные примеры из области физики можно было бы продолжить, но надо иметь в виду, что влияние науки на развитие техники не является монополией физики. Не меньшее значение имеют и другие отрасли науки.

Развитие химической науки оказало существеннейшее влияние на многие производства. Здесь в первую очередь необходимо отметить все возрастающее значение производства полимеров. Трудно найти такую отрасль промышленности, которая не является потребителем полимерных материалов. В машиностроении наблюдается тенденция к замене металлов полимерами при изготовлении ряда изделий и деталей машин. Полимерные пленки находят применение в строительстве и сельском хозяйстве. Строятся экспериментальные дома, которые почти целиком сооружены из полимерных материалов.

Огромна роль химии в повышении продуктивности сельского хозяйства. Важнейшими направлениями здесь являются: созда-

ние различных типов удобрений, защита растений от вредителей, применение искусственных кормовых средств в животноводстве, развитие методов и средств консервирования сельскохозяйственной продукции, исследование состава и свойств почв и другие направления. Особо следует отметить синтез пищевых продуктов из непищевого сырья: производство спирта и белков из нефти, производство искусственных жиров. Эти работы позволяют направить для непосредственного употребления в пищу часть растительного и животного сырья, которое раньше использовалось в технических целях.

В перспективе эти работы позволят расширить наши пищевые ресурсы синтетическими продуктами.

Велики достижения биологической науки и их воздействие на сельскохозяйственную практику и медицину. Изучение законов передачи наследственной информации дает в руки практиков-селекционеров мощные средства направленного изменения свойств организмов. Биология в содружестве с химией дает возможность значительно увеличить производство сельскохозяйственной продукции. Здесь следует отметить применение химических мутагенов, вызывающих полезные наследственные мутации у растений и полиплоидию — явление кратного увеличения набора хромосом в клетках растений. Полиплоидия сопровождается изменением многих свойств растения, в частности увеличением размеров плодов и содержания в них полезных веществ: белков, сахара и других. На достижения биологической науки в последние годы все больше и больше опирается медицинская практика.

Особенно наглядными примерами большого значения науки для создания новых технических средств являются работы по освоению космического пространства. Создание космических кораблей, ракет для их запуска, оборудования для обеспечения жизнедеятельности человека в космосе стало возможным благодаря огромному числу предварительных научных исследований в области физики, химии, металлургии, радиоэлектроники и многих других отраслей науки. В космическом корабле, как в фокусе, сконцентрирована научная и техническая мысль исследователей, конструкторов, инженеров и новаторов производства.

«Сейчас во всем мире происходит чрезвычайное убыстрение использования достижений науки. Если в прежние времена наука и техника развивались в значительной мере обособленно и достижения науки лишь постепенно влияли на технический прогресс, то сейчас общество в гораздо большей степени развивает научные исследования направленно, с целью открыть новые источники технического прогресса. . .

Процесс сближения науки и промышленности, их взаимного оплодотворения привел к громадному росту масштабов научных исследований, к невиданному ускорению развития науки и техники, открытию возможностей резкого увеличения объемов

производства. Этот процесс характеризуют как современную научно-техническую революцию. Значение науки в развитии нашего общества, в строительстве коммунизма будет неуклонно возрастать» [4].

В настоящее время наука достигла такого уровня развития, когда производительность труда ученых и инженеров все в большей степени зависит от правильной организации научного труда, от вооруженности исследований новейшей техникой и от эффективности научно-информационного обслуживания исследователя, постановки системы научной информации. При этом необходимо подчеркнуть, что научно-информационная деятельность — специфическая разновидность научной работы, возникшая в результате общественного разделения труда в науке. В научно-информационной деятельности существует своя проблематика и применяются свои методы решения стоящих перед нею задач. Сейчас эволюция научно-информационной деятельности достигла такого этапа, когда на основе этой деятельности начала складываться новая научная дисциплина — информатика.

Формирование информатики как самостоятельной научной дисциплины происходит в результате действия двух факторов:

а) внутренних закономерностей развития самой науки;

б) потребностей общественной практики.

## ЗАКОНОМЕРНОСТИ РАЗВИТИЯ НАУКИ И НАУЧНАЯ ИНФОРМАЦИЯ

Одной из основных закономерностей развития науки является ее преемственность и интернациональный характер. Наука строит свое здание на фундаменте знаний, накопленных предыдущими поколениями. Если бы каждому поколению нужно было заново открывать все основные законы природы и общества, то наука развивалась бы очень медленно. Поэтому К. Маркс назвал науку продуктом всеобщего исторического процесса развития. «Всеобщим трудом является всякий научный труд, всякое открытие, всякое изобретение. Он обусловливается частью кооперацией современников, частью использованием труда предшественников» [5].

Преемственность — важный фактор ускорения развития науки. Последующие поколения овладевают всей совокупностью научных знаний, накопленных их предшественниками, в основном не путем изучения первоисточников. Для этого используются материалы, в которых научные данные излагаются в обработанном и систематизированном виде. Поэтому для ускорения развития науки каждое поколение ученых должно не только добывать новые

---

[4] М. В. Келдыш. Речь на XXIII съезде КПСС. «Правда», 3 апреля 1966 г.

[5] К. Маркс. Капитал, т. 3. М., Госполитиздат, 1955, стр. 109.

научные данные, но и проводить огромную работу по систематизации всей суммы научных знаний своего и всех предшествующих поколений.

До последнего времени эта исключительно важная работа проводилась лишь на базе имеющегося опыта и личной инициативы. Однако в наши дни объем накопленной человечеством научной информации достиг столь огромных размеров, а количество добываемых учеными новых научных данных так велико, что традиционные методы обработки и систематизации научной информации становятся все более недостаточными. Для успешного проведения такой работы необходима разработка ее научных основ. Это позволило бы широко применить для обработки и систематизации научной информации средства современной техники, в том числе различные информационные машины, и таким образом ускорить общее развитие науки и техники.

Преемственность в развитии науки неразрывно связана с ее интернациональным характером. Большинство научных открытий и изобретений было сделано в результате труда ученых, работающих в разных странах мира. «Критическая история технологии, — писал К. Маркс в «Капитале», — вообще показала бы, как мало какое бы то ни было изобретение XVIII столетия принадлежит тому или иному отдельному лицу» [6].

174028

Справедливость этого положения отчетливо видна хотя бы на примере создания кибернетики — науки об управлении сложными динамическими системами. Основным создателем кибернетики вполне заслуженно считается американский математик Норберт Винер. Однако в своих исследованиях он широко использовал результаты работ многих других ученых, особенно И. П. Павлова и А. Н. Колмогорова (СССР), А. Тьюринга и Г. Уолтера (Великобритания), К. Шеннона и Дж. Байглоу (США), А. Розенблюта (Мексика) и др. Характеризуя вклад, который внес в создание кибернетики академик А. Н. Колмогоров, опубликовавший незадолго до второй мировой войны в «Докладах Французской академии наук» статью о прогнозировании дискретных последовательностей, Норберт Винер писал: «. . . Вполне возможно утверждать, что Колмогоров не только независимо разобрал все основные вопросы этой области, но и был первым, опубликовавшим свои результаты» [7].

Основным средством передачи научных знаний во времени и пространстве являются письменные документы — книги, журнальные статьи и другие публикации, в которых ученые излагают результаты своих исследований. Наука возникает лишь с появлением письменности, так как только письменность может обеспечить подлинную преемственность в развитии науки и только

[6] К. Маркс и Ф. Энгельс. Сочинения, т. 23, стр. 383.
[7] Н. Винер. Я — математик. М., «Наука», 1964, стр. 250.



БИБЛИОТЕКА

благодаря письменности наука приобретает подлинно интернациональный характер.

Для того чтобы научные документы наилучшим образом выполняли свое основное назначение — служить средством передачи научных знаний из одного географического пункта в другой, от одного поколения людей к другому, необходима разработка научных принципов подготовки различных видов документов. Такие принципы могут быть разработаны лишь на основе глубокого изучения психологии творческого мышления исследователей, а также особенностей восприятия и усвоения ими письменной информации. В настоящее время необходимость этого начинает осознавать все большее число ученых.

Многие видные ученые и специалисты прошлого хорошо понимали, что на них лежит большая ответственность за распространение результатов проводимых ими исследований. Поэтому они уделяли много сил и внимания подготовке рукописей своих книг и статей, справедливо считая это органической частью научной работы. Великий английский физик Э. Резерфорд вообще не считал научную работу завершенной, пока она не описана и не опубликована [8]. Работавший у него академик Ю. Б. Харитон вспоминает, что Э. Резерфорд «предъявлял очень строгие требования к изложению результатов, часто заставляя полностью переделывать уже написанные статьи» [9].

Другой важной закономерностью развития науки является взаимодействие ее отдельных отраслей, на основе которого возникают и развиваются новые науки. Такое взаимодействие происходит не только между смежными, но и между далеко отстоящими друг от друга науками. Это приводит к образованию множества стыков между самыми различными науками пограничных областей знания. На таких стыках находятся наиболее активные точки роста современной науки, именно там делаются наиболее важные научные открытия. Эту особенность развития науки отметил еще Ф. Энгельс. В «Диалектике природы» он писал, что именно в местах соприкосновения наук *надо ожидать наибольших результатов* [10].

Особенно сильное взаимодействие наук происходит на современном этапе их развития. В результате такого взаимодействия возникли и быстро развиваются такие важные науки и научные дисциплины, как биохимия и биофизика, физическая химия, гео-

---

[8] A. S. Eve. Rutherford, being the life and letters of the Rt Hon. Lord Rutherford. Cambridge, University Press, 1935, стр. 406. Цит. по: «Journal of Documentation», 1964, v. 20, № 1, стр. 2.

[9] Цит. по: И. Б. Файнбойм. Эрнест Резерфорд. М., «Знание», 1958, стр. 40.

[10] Ф. Энгельс. Диалектика природы. М., Госполитиздат, 1955, стр. 235.

химия и биогеохимия, космическая медицина, бионика, кибернетика, семиотика, инженерная психология и многие другие.

Взаимодействие наук и их синтез происходят в условиях постоянной и все более усиливающейся дифференциации и специализации наук. В настоящее время общая сумма знаний, накопленных даже в какой-либо одной отрасли науки, растет значительно быстрей возможности их усвоения каждым отдельным ученым. Поэтому ученые вынуждены сосредоточивать свои усилия на изучении все более узких разделов той или иной науки, что неизбежно ведет к ее делению на самостоятельные научные дисциплины.

«Любая область науки, — отмечал академик А. Н. Несмеянов, — дифференцируется и порождает все новые и новые ветви. Так, химия уже в прошлом веке разделилась на органическую и неорганическую, аналитическую и затем физическую химию. В нынешнем столетии появилась физическая органическая химия, а сравнительно недавно и элементоорганическая химия. В физике выделились в качестве самостоятельных ее ветвей с собственными методами исследований молекулярная физика, оптика, акустика, физика твердого тела, в последнее десятилетие — ядерная физика и т. д.» [11].

Прогресс науки заключается в переходе от раскрытия простых причинно-следственных отношений и связей к формулированию все более глубоких и фундаментальных законов природы и общества. Исследователи добывают все новые и новые научные факты. На основе изучения и обобщения этих фактов они открывают новые законы и создают новые теории, заменяя ими теории устаревшие. Этот процесс не имеет конца, поскольку материя обладает бесконечным разнообразием свойств, бесконечным числом видов, форм движения и их взаимосвязей.

Внутренняя логика научного исследования такова, что чем больше ученый узнает о предмете своего исследования, тем больше новых аспектов этого предмета и его связей с другими предметами он выявляет, тем шире раздвигается перед ним фронт исследований. В течение некоторого времени ученый еще пытается вести исследования по всему этому фронту. Но неизбежно наступает такой момент, когда это становится ему не под силу. И тогда ученый выбирает какой-то узкий участок этого фронта и концентрирует все свои силы на этом участке, превращаясь в еще более узкого специалиста. Лишь при таком условии он может рассчитывать на успех в своей работе. Таким образом, дифференциация и специализация науки есть следствие и в то же время непременное условие ее развития.

---

[11] А. Н. Несмеянов. На стыке наук. «Наука и жизнь», 1958, № 3, стр. 1.

Дифференциация и специализация науки — это объективно действующая закономерность. Она помогает человечеству все глубже проникать в тайны мироздания, и в этом ее несомненная положительная сторона. Но она имеет и свою отрицательную сторону. Неизбежным следствием дифференциации и специализации науки является рост обособленности ученых. Представителям разных научных специальностей даже в пределах одной науки становится все трудней понимать друг друга. Кроме того, узкие специалисты начинают терять общую ориентировку во всей системе научных знаний и в разнообразных явлениях природы. И это происходит в то время, когда взаимодействие между науками стало играть все более важную роль в их развитии. В таких условиях дальнейшее углубление дифференциации и специализации науки начинает сильно мешать ее успешному развитию.

Эта исключительно важная и трудная проблема частично решается взаимопроникновением наук, их интеграцией. Но основную роль в ее решении призвана сыграть информатика, которая должна создать надежные каналы связи между разными науками, обеспечить систематический обмен научными данными между ними и снабдить ученых всех специальностей эффективными средствами ориентировки в расширяющейся и усложняющейся системе научных знаний. «Прогресс науки, — подчеркивает академик А. Н. Несмеянов, — прямо связан с живым обсуждением стоящих перед наукой проблем, с хорошей организацией научной информации, дающей ученым возможность следить как за работами в области своей специальности, так и за сопредельными науками» [12].

Рассматривая перспективы развития науки в ближайшие 100 лет и его возможные последствия, известный английский ученый — физик Джордж Томсон писал в своей книге «Предвидимое будущее»: «Совершенно неизбежно, что по мере расширения наших знаний та их доля, которой в состоянии овладеть один человек, будет убывать. Это обстоятельство влечет за собой по меньшей мере два нежелательных последствия. Поскольку для работы по избранной профессии человеку требуется знать очень многое, он испытывает величайший соблазн учить как можно меньше из того, что ему в работе непосредственно не пригодится. У него создается однобокое представление о мире, в котором он живет. . .

Второе из нежелательных последствий . . . проявляется на более поздней стадии специализации. Многие достижения науки и техники являются следствием внедрения тех или иных идей и технических приемов в областях, для которых они первоначально не предназначались. Однако реализация подобной возможности требует все труднее и труднее достижимой широты знаний. Вполне

[12] А. Н. Несмеянов. На стыке наук. «Наука и жизнь», 1958, № 3, стр. 2.

может случиться, что эта особенность воздвигает предел прогрессу науки, но здесь многое можно сделать для того, чтобы отодвинуть приход этой катастрофы. В частности, мы только сейчас начали систематически делать новые открытия доступными специалистам в смежных областях науки. Эти попытки популяризации нового нужно значительно расширять. Это — нелегкая задача и те, кому это удается, заслуживают того, чтобы их работа ценилась научной общественностью так же высоко, как и работа исследователей. . . По мере раздвижения границ науки необходимо будет, помимо подобных популяризаторов, предусмотреть еще и специальных «связных» между отдельными отраслями науки, в обязанности которых входило бы обращать внимание ученых на те моменты, по которым возможно установление полезных связей между этими отраслями. В науке неминуемо должно будет произойти то, что произошло в армии. В целях обеспечения боеспособности солдата на передовой линии приходится в тылу ставить за ним все больше и больше людей. То же самое наблюдается сейчас и с людьми, занимающимися научно-исследовательской работой» [13].

Важным законом развития науки является закон ее ускоренного развития. Честь открытия этого закона принадлежит Ф. Энгельсу. В статье «Наброски к критике политической экономии», опубликованной в 1844 г., он писал: «. . .наука движется вперед пропорционально массе знаний, унаследованных ею от предшествующего поколения» [14].

Закон ускоренного развития науки — есть частный случай открытого также Ф. Энгельсом всеобщего закона ускоренного развития живой природы и общества. «По отношению ко всей истории развития органов, — писал Ф. Энгельс, — надо принять закон ускорения пропорционально квадрату расстояния во времени от исходного пункта». В подтверждение этого Ф. Энгельс ссылается на соответствующие работы Э. Геккеля и делает вывод: «Чем выше, тем быстрее идет дело» [15]. Все данные современной науки свидетельствуют о том, что этот основной вывод Ф. Энгельса полностью оправдывается. Однако вопрос о форме восходящей кривой требует уточнения. Не исключена возможность, что предложенная Ф. Энгельсом кривая параболического типа (пропорциональность квадрату времени) окажется неточной [16].

В области естественных наук закон ускоренного развития начал действовать с момента открытия Коперником гелиоцентрической системы. Указывая на это, Ф. Энгельс писал: «Чем в рели-

---

[13] Дж. Томсон. Предвидимое будущее. М., ИЛ, 1958, стр. 158—159.

[14] К. Маркс и Ф. Энгельс. Сочинения, т. 1, стр. 568.

[15] Ф. Энгельс. Диалектика природы. М., Госполитиздат, 1955, стр. 247.

[16] М. Н. Руткевич. Диалектический материализм. М., Соцэкгиз, 1959, стр. 392—394.

гиозной области было сожжение Лютером папской буллы, тем в естествознании было великое творение Коперника, в котором он, — хотя и робко, после 36-летних колебаний и, так сказать, на смертном одре, — бросил вызов церковному суеверию. С этого времени исследование природы по существу освободилось от религии, хотя окончательное выяснение всех подробностей затянулось до настоящего времени и далеко еще не завершилось во многих головах. Но с тех пор и развитие науки пошло гигантскими шагами, ускоряясь, так сказать, пропорционально квадртату удаления во времени от своего исходного пункта, как бы желая показать миру, что по отношению к движению высшего цвета органической материи, человеческому духу, имеет силу закон, обратный закону движения неорганической материи »[17].

Ускорение развития науки происходит под воздействием многих факторов. Основными из них можно считать:

1) непрерывное накопление научных знаний, обусловленное преемственностью в науке и ее интернациональным характером;

2) углубляющуюся дифференциацию и специализацию науки и связанное с ними повышение производительности научного труда;

3) усиление взаимодействия наук;

4) экономический, социальный и политический строй общества, содействующий прогрессу науки или тормозящий ее развитие;

5) ускорение развития производительных сил общества;

6) совершенствование средств сообщения и связи;

7) увеличение средств, расходуемых на научные исследования;

8) увеличение числа лиц, занимающихся научными исследованиями.

Сформулированный Ф. Энгельсом закон ускоренного развития науки гласит: «Наука движется вперед пропорционально массе знаний, унаследованных ею от предшествующего поколения. . .» На основе этого положения может быть составлено дифференциальное уравнение

$$\frac{dy}{dt} = ky,$$

проинтегрировав которое, мы получим математическое выражение закона ускоренного развития науки:

$$y = ae^{kt} \, (k > 0),$$

где $y$ — сумма знаний, унаследованная от прошлых поколений; $a$ и $k$ — постоянные величины; $e$ — основание натуральных логарифмов ($e$=2,718. . .), $t$ — время.

---

[17] Ф. Энгельс. Диалектика природы. М., Госполитиздат, 1955, стр. 153.

Такая зависимость в математике называется экспоненциальной функцией.

Смысл величины $a$ легко определить, приняв $t=0$. Тогда $y=a$, откуда следует, что $a$ — сумма научных знаний к началу отсчета. Коэффициент $k$, видимо, не является постоянным для всех эпох и стран и зависит от уровня развития промышленности и других условий [18].

Выше были перечислены основные факторы, от которых зависят темпы развития науки. Два последних фактора — размеры средств, расходуемых на научные исследования, и численность научных работников — имеют количественное выражение. Очевидно, что в настоящее время темпы развития науки находятся в прямой зависимости от этих двух факторов. Поэтому справедливость закона об ускоренном (экспоненциальном) развитии науки может быть проверена на статистических данных, характеризующих динамику роста ассигнований на научно-исследовательские работы, а также динамику роста численности научных работников.

На рис. 1—3 приведены кривые роста ассигнований на научно-исследовательские и опытно-конструкторские работы в США. Из этих рисунков видно, что в течение последних 40 лет эти расходы увеличиваются экспоненциально, возрастая примерно на 10—15% в год, и что из всех расходов на научно-исследовательские и опытно-конструкторские работы, которые были сделаны в США за всю историю этого государства, около 50% приходится на 5 лет с 1961 по 1965 г. В связи с этим нельзя не отметить, что в последние годы основная часть этих расходов идет на расширение гонки вооружений, которую развязало правительство США.

На рис. 4 приведена кривая роста числа статей в последовательных изданиях справочника «American Men of Science», вышедших за последние 50 лет. Можно сделать вполне разумное предположение, что число таких статей в справочнике увеличивалось в прямой зависимости от роста численности американских ученых. Тогда на рис. 4 следует, что в течение последних 50 лет численность американских ученых росла по экспоненте и удваивалась примерно каждые 12 лет. На рис. 5 приведена кривая роста численности научных работников в СССР, из которой видно, что этот рост также происходит по экспоненте, причем численность советских ученых удваивается каждые 6—8 лет.

Одним из важных следствий ускоренного развития науки является экспоненциальный рост количества выходящей в мире научно-технической литературы. На рис. 6 показаны кривые роста числа научно-технических и реферативных журналов за последние 300 лет (данные Д. Прайса) [19]. Эти кривые в целом

[18] М. М. Карпов. Закон ускоренного развития естественных наук. «Вопросы философии», 1963, № 4, стр. 108—111.

[19] D. J. Price. The exponential curve of science. «Discovery», 1956, 17, № 6, стр. 240—243.

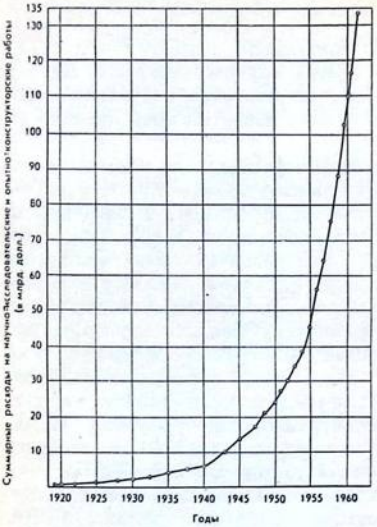



Рис. 1. Суммарные расходы на научно-исследовательские и опытно-конструкторские работы в США

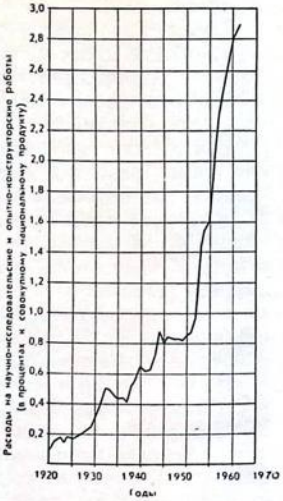



Рис. 3. Рост доли расходов на научно-исследовательские и опытно-конструкторские работы по отношению к национальному продукту США

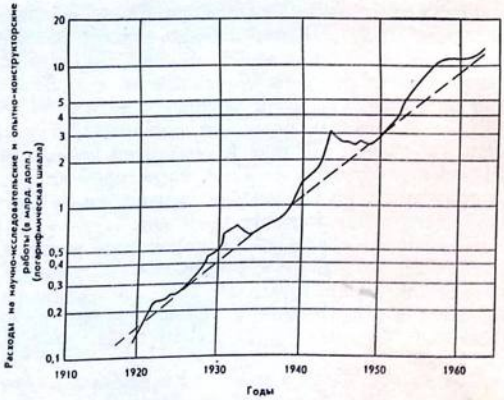



Рис. 2. Расходы на научно-исследовательские и опытно-конструкторские работы в США (по годам)

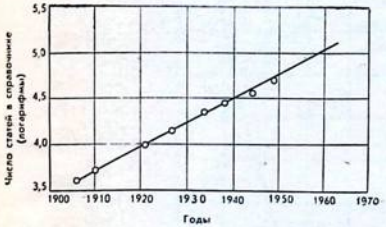



Рис. 4. Рост числа статей в справочнике «American Men of Science»

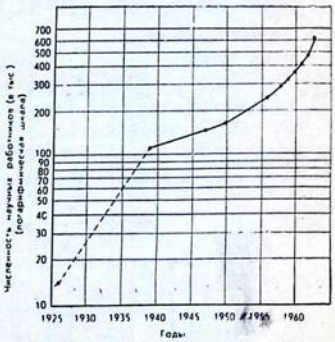



Рис. 5. Рост численности научных работников в СССР

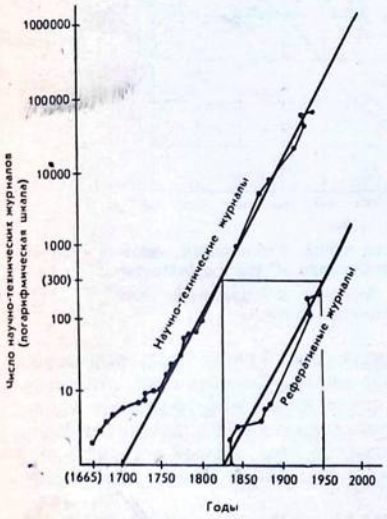



Рис. 6. Рост числа выходящих в мире научно-технических и реферативных журналов

подтверждают, что в течение указанного периода рост числа научно-технических и реферативных журналов происходил по экспоненте (кривые построены в полулогарифмических координатах). Экстраполяция этих кривых в историческое прошлое приводит

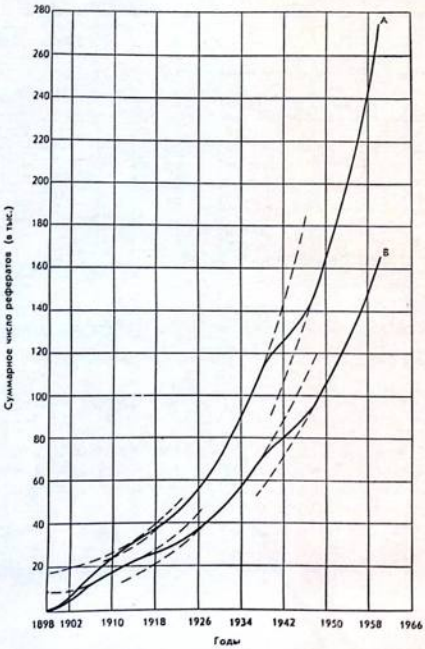



**Рис. 7. Рост суммарного числа публикаций, освещенных в реферативном журнале «Science Abstracts»**

*A* — для серии «Physics Abstracts»; *B* — для серии «Electrical Engineering Abstracts»

к разумным результатам. Ордината кривой роста числа научно-технических журналов получает значение, равное единице, примерно в 1700 г., а ордината кривой роста числа реферативных журналов — примерно в 1825 г. Это довольно точно соответствует фактическому положению вещей. Известно, что первый научный журнал «Le journal des scavants» появился в 1665 г.[20], а первый рефера-

[20] Некоторые исследователи считают первым научным журналом «Philosophical Transactions», который начал издаваться Королевским обществом Англии также с 1665 г.

тивный журнал «Pharmaceutisches Centralblatt» вышел из печати в 1830 г.

Об экспоненциальном росте количества научно-технической литературы свидетельствует также непрерывное увеличение числа публикаций, освещаемых в ведущих реферативных журналах мира. На рис. 7 показаны кривые роста суммарного числа публикаций, освещенных в реферативном журнале «Science Abstracts» [21]. Из этого рисунка видно, что нарушения экспоненциального хода кри-

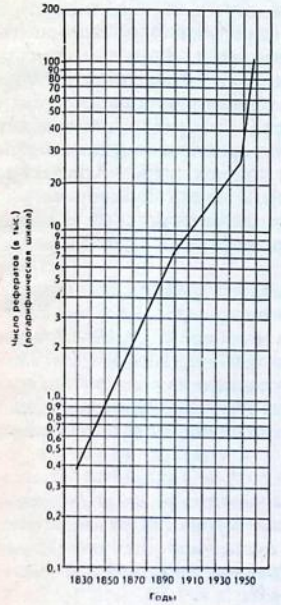



Рис. 8. Рост числа публикаций, ежегодно освещаемых в реферативном журнале «Chemisches Zentralblatt»

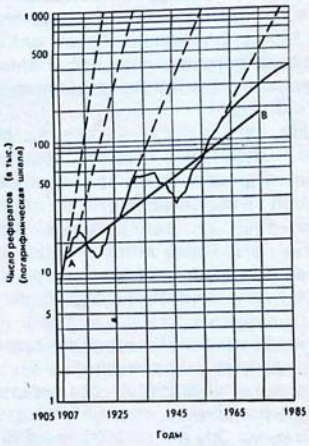



Рис. 9. Кривая роста публикаций, ежегодно освещаемых в реферативном журнале «Chemical Abstracts»

вых наблюдались только во время первой и второй мировых войн. После окончания этих войн рост суммарного числа рефератов, помещенных в журнале «Science Abstracts», продолжался синбатно первоначальным кривым.

Примерно такой же характер имеют кривые роста числа публикаций, ежегодно освещаемых в реферативных журналах по хи-

[21] Реферативный журнал «Science Abstracts» выходит в Великобритании с 1898 г. в двух сериях: «Physics Abstracts» и «Electrical Engineering Abstracts» («Electrical and Electronic Engineering Abstracts»).

мии «Chemisches Zentralblatt» (выходит в Берлине с 1830 г.) и «Chemical Abstracts» (выходит в США с 1907 г.). См. рис. 8 и 9.

За 59 лет существования реферативного журнала «Chemical Abstracts» (т. е. с 1907 по 1966 г.) в нем было помещено 4 млн. рефератов, из которых свыше 0,9 млн. (26%) было опубликовано за последние 5 лет. В период с 1951 по 1965 г. ежегодное увеличение числа рефератов, помещаемых в журнале «Chemical Abstracts», составляло в среднем 8,6%. По формуле сложных процентов можно легко подсчитать, что число рефератов, ежегодно помещаемых в журнале «Chemical Abstracts» (а следовательно, и объем литературы по химии), в настоящее время удваивается каждые 8,2 года. Удвоение суммарного числа рефератов, помещенных в журнале «Chemical Abstracts» за все время его существования (т. е. с 1907 г.), в настоящее время происходит каждые 11 лет (до 1950 г. оно происходило за 17 лет). При такой скорости роста числа рефератов, ежегодно помещаемых в журнале «Chemical Abstracts», к 1971 г. суммарное число рефератов должно достигнуть 5 млн.[22]

На рис. 9 показана кривая роста числа рефератов, помещенных в журнале «Chemical Abstracts» за период с 1907 по 1960 г. Кривая дана в полулогарифмических координатах. Анализ этой кривой показывает, что в периоды 1907—1913, 1918—1930 и 1945—1960 гг. она довольно точно совпадала с экспонентой. Однако в каждый последующий период крутизна этой экспоненты уменьшается, и экспонента проявляет тенденцию к постепенному переходу в линейную функцию.

Примерно такая же картина наблюдается в области биологии. За последние 200 лет количество выходящих в мире журналов по биологии (в их число не включены журналы по медицине и сельскому хозяйству) удваивалось в среднем каждые 18 лет. Следует отметить, что 65% этих журналов появилось в течение последних 30 лет (с 1927 по 1957 г.) [23].

При рассмотрении темпов количественного роста научно-технической литературы не следует упускать из виду, что хотя этот рост происходит по экспоненте, крутизна такой экспоненты постепенно уменьшается. Поэтому необходимо с некоторой осторожностью использовать экспоненциальные кривые с постоянным $k$ для прогнозирования объема, который будет иметь научно-техническая литература в будущем. Как показывает практиче-

---

[22] D. B. Baker. Growth of chemical literature. Past, present and future. «Chemical and Engineering News», 1961, v. 39, № 29, стр. 78—81; Chemical literature expands. «Chemical and Engineering News», 1966, v. 44, № 23, стр. 84—88.

[23] M. G. Conrad. Growth of biological literature and the future of «Biological Abstracts». «Federation Proceedings», 1957, v. 16, № 3, стр. 711—712.

ский опыт, трудно ожидать надежных результатов, если экстраполяция таких кривых производится в будущее более чем на 10 лет.

Необходимо также учитывать, что на журнальные статьи приходится лишь около $^1/_3$ всех научно-технических публикаций, и $^2/_3$ составляют научно-технические отчеты и подобные им документы. Удельный вес таких публикаций в научно-технической литературе быстро растет, но этот рост очень трудно учитывать. Нельзя также забывать о все более глубокой милитаризации науки в странах капитализма, об обострении конкурентной борьбы между промышленными фирмами и о расширении масштабов промышленного шпионажа, что ведет ко все большему ограничению публикаций результатов научно-исследовательских и опытно-конструкторских работ. Проф. Д. Бернал отмечал, что «в области физических наук в Англии и Америке около двух третей информации о новых открытиях нигде не публикуется, а находится в архивах фирм и тщательно охраняется от конкурентов, которые зачастую открывают то же самое и скрывают свои открытия» [24].

Если считать, что основными показателями развития науки являются размеры ассигнований на научные исследования, количество научных работников и число научных публикаций и что рост этих показателей и в будущем будет происходить по экспоненте, то должно наступить такое время, когда число научных публикаций и количество научных работников сравняется с численностью народонаселения, а расходы на научные исследования составят 100% национального продукта той или иной страны. Но это явный абсурд.

Известный американский историк науки Д. Прайс, анализируя эту парадоксальную ситуацию, приходит к выводу, что темп роста количественных показателей развития науки должен замедлиться. В своих рассуждениях он исходит из того, что в природе экспоненциальный рост всегда приводит к насыщению. Но эту справедливую для живой природы закономерность Д. Прайс, допуская типичную для большинства буржуазных ученых ошибку, произвольно переносит на человеческое общество и делает вывод, что развитие науки должно замедлиться и, достигнув состояния насыщения, совсем прекратиться.

В своей книге «Малая наука, большая наука» он пишет: «Принимая за доказанное, что, когда мы говорим о науке, речь идет о закономерном экспоненциальном росте с периодом удвоения от 10 до 15 лет, мы можем предположить, что существование «потолка» — вещь весьма правдоподобная, поскольку в ином случае мы получили бы абсурдные условия к концу столетнего периода. Но если принято существование такого предела, мы

---

[24] Д. Бернал. Успех науки. «Вопросы философии», 1959, № 8, стр. 61.

должны будем прийти к выводу, что экспоненциальный рост науки есть лишь начало логистической кривой в другом обличье. . . Таким образом, даже без ссылок на современное положение дел или на оценки того, когда и где объявится «потолок», ограничивающий рост, для нас становится очевидным, что в пределах жизни одного человеческого поколения наука должна будет отказаться от традиционного экспоненциального роста и приблизиться к критической точке, маркирующей ее переход к старческой дряхлости» [25].

Еще одной важной причиной, по которой Д. Прайс приходит к глубоко ошибочной теории предела развития науки, является антиисторичность его мышления. Как справедливо отметил В. Н. Столетов в послесловии к сборнику «Наука о науке», в котором опубликован перевод на русский язык цитированной выше книги Д. Прайса, последний исходит из неизменности всех условий развития науки: экономической структуры мира, социального строя, уровня образования, доступности научных знаний для народа и т. д. Будучи историком науки, Д. Прайс пользуется абстрактно-историческим методом. Он не учитывает диалектического положения, по которому количественный рост неизбежно приводит к качественным изменениям. За внешним сходством темпов роста показателей науки теперь, сто и двести лет назад, он не видит, что эти явления в настоящее время приобретают принципиально иное значение. Естественно, что рост численности ученых и размеров ассигнований на научные исследования не может все время продолжаться с такой же скоростью, что и в настоящее время. Но замедление темпов роста этих показателей вовсе не будет ни причиной, ни свидетельством замедления темпов развития науки. Уже в настоящее время начинают действовать факторы, которые обеспечивают возможность неограниченного и все более растущего темпа развития науки без экспоненциального роста численности ученых и расходов на научные исследования. Пока основными из этих факторов являются:

1) совершенствование организации и методики проведения научных исследований;

2) оснащение научно-исследовательских работ современной техникой;

3) широкое применение в научной работе различных кибернетических средств, моделирующих отдельные процессы мышления;

4) создание эффективных информационных систем, позволяющих ученым значительно уменьшить расходы рабочего времени на информационную работу;

5) сокращение сроков подготовки научных работников (путем оптимизации методов обучения).

---

[25] Д. Прайс. Малая наука, большая наука. В кн.: «Наука о науке». М., 1966, стр. 303.

Первые четыре фактора обеспечивают непрерывное повышение производительности научного труда исследователей, а пятый фактор увеличивает продолжительность их творческой жизни. Все это должно компенсировать соответствующее снижение темпов роста численности ученых и размеров ассигнований на научные исследования и в то же время обеспечить дальнейшее ускорение развития науки и техники.

Что же касается проблемы экспоненциального роста количества научных публикаций, то она может иметь несколько решений. Уже сейчас результаты научных исследований далеко не всегда сообщаются в опубликованных работах. Не может быть сомнения, что в недалеком будущем на смену научным публикациям в их традиционной форме начнут приходить и другие, более совершенные методы и средства передачи и распространения документальной научной информации. По мере их внедрения вместо этой проблемы будут возникать новые, связанные со сбором документальной информации, ее обработкой и созданием механизированных и автоматизированных информационно-поисковых систем.

## ПОТРЕБНОСТИ ОБЩЕСТВЕННОЙ ПРАКТИКИ И НАУЧНАЯ ИНФОРМАЦИЯ

Непрерывное увеличение средств, расходуемых на научно-исследовательские и опытно-конструкторские работы, а также вовлечение все большего числа людей в сферу научного труда являются важными показателями роста науки и играют все бóльшую роль в развитии современного общества. В настоящее время наиболее развитые в промышленном отношении страны расходуют на эти цели до 3% своего национального дохода, и с каждым годом эта цифра увеличивается. В 1961 г. в США на научно-исследовательские и опытно-конструкторские работы было израсходовано 2,84% национального дохода [26], а в Великобритании в в 1962 г. — около 2,5% [27].

Непрерывно растет и процент населения, занятого в сфере науки и техники. Для США это подтверждается данными [28] табл. 1.

В настоящее время научными исследованиями в мире занято более 2 млн. человек, что составляет 90% общей численности

---

[26] National Science Foundation. 13th Annual Report. NSF—64.1, 1963, стр. 132—133.

[27] H. Coblans. Science, government, and information. «Journal of Documentation», 1964, v. 20, № 1, стр. 1.

[28] National Science Foundation. 13th Annual Report. NSF—64.1, 1963, стр. 132. Engineering and scientific manpower in the United States, Western Europe and Soviet Russia. Joint Committee on Atomic Energy. 84th Congress, 2nd Session. Washington, U. S. Government Printing Office, 1956, стр. 6—7.

Таблица 1

**Рост занятости в сфере науки и техники в США по годам**
(млн. человек)

| Показатели | 1920 | 1930 | 1940 | 1950 | 1960 | 1965 |
|---|---|---|---|---|---|---|
| Население США . . . . . | 105,7 | 123,1 | 132,0 | 152,3 | 180,7 | 209 |
| Общая численность занятого населения . . . . . | 42,4 | 48,6 | 56,2 | 64,7 | 73,1 | 76 |
| Численность занятых в сфере науки и техники . | | | | | | |
| в абсолютных цифрах . . . . . | 0,29 | 0,45 | 0,86 | 1,47 | 2,37 | 2,7 |
| в процентах . . . . . | 0,7 | 0,9 | 1,5 | 2,2 | 3,2 | 3,6 |

ученых и инженеров за всю историю человечества [29]. В Советском Союзе в 1966 г. насчитывалось 711,6 тыс. научных работников [30]. Академик Н. Н. Семенов считает, что в будущем половина человечества тем или иным путем будет участвовать в созидательном научном труде [31].

В таких условиях исключительное значение приобретает повышение производительности научного труда, которое становится все более суровым требованием общественной практики. Превращение науки в решающий фактор дальнейшего развития производительных сил социалистического общества, вступившего на путь построения коммунизма, выдвигает на передний план проблему повышения производительности научного труда.

«Капитализм, — писал В. И. Ленин, — может быть окончательно побежден и будет окончательно побежден тем, что социализм создает новую, гораздо более высокую производительность труда» [32].

Повышение производительности научного труда все больше происходит благодаря специализации научной работы, ее общественному разделению. Известно, что характер и формы разделения труда определяются развитием производительных сил общества и что разделение труда вызывает дальнейшее повышение производительности труда. В настоящее время уже близится к завершению процесс разделения научного труда на теоретические и экспериментальные исследования. «Из истории физики, —

[29] П. Ожe. Современные тенденции в научных исследованиях. М., ЮНЕСКО, 1963, стр. 5.

[30] СССР в цифрах в 1966 г. Краткий статистический сборник. М., «Статистика», 1967, стр. 160.

[31] Цит. по: П. Л. Капица. Будущее науки. В кн.: «Наука о науке». М., 1966, стр. 111.

[32] В. И. Ленин. Полное собрание сочинений, т. 39, стр. 21.

писал академик П. Л. Капица, — хорошо известно, что деление физиков на теоретиков и экспериментаторов произошло совсем недавно. В прежние времена не только Ньютон и Гюйгенс, но и такие теортики, как Максвелл, обычно сами экспериментально проверяли свои теоретические выводы и построения. Теперь же только в исключительных случаях теоретик ставит опыты, чтобы проверить свои теории» [33].

Разделение исследователей на теоретиков и экспериментаторов объясняется применением в исследованиях все более сложного математического аппарата, а также значительным усложнением техники эксперимента. В настоящее время эксперимент становится не под силу одному человеку и обычно выполняется коллективом научных работников, которые обладают специальной подготовкой и используют сложное и дорогостоящее оборудование.

Третьей разновидностью научного труда все в большей степени становится научно-информационная деятельность. Раньше ученый-исследователь мог без особого труда следить за соответствующей научной литературой, а также непосредственно обмениваться соответствующей информацией со своими коллегами и, таким образом, быть в курсе последних достижений науки и техники. Сейчас положение коренным образом изменилось. Несмотря на то, что современные достижения науки и техники в огромной мере облегчили возможности личных контактов, сам факт вовлечения в научную деятельность большого числа людей сделал такие контакты весьма затруднительными. В настоящее время над решением одной научной проблемы работают десятки, а нередко и сотни научных работников.

Важным фактором, делающим для исследователя практически невозможным самостоятельное ознакомление с интересующей его научно-технической литературой, является огромное количество этой литературы. Об этом могут свидетельствовать следующие цифры. По данным Библиотеки конгресса США, сейчас в мире выходит около 30 тыс. научно-технических журналов, в которых ежегодно публикуется от 0,9 до 2,1 млн. статей [34]. В других источниках указывается, что в настоящее время число научно-технических журналов достигло уже 50 тыс. и продолжает расти со скоростью 10—15% в год [35]. Быстро увеличивается также число научно-технических отчетов и других изданий подобного типа. Только в США ежегодно публикуется 100—150 тыс. научно-тех-

---

[33] П. Капица. Теория, эксперимент, практика. М., «Знание», 1966, стр. 16.

[34] C. P. Bourne. Methods of information handling. New York, Wiley, 1963, стр. 2.

[35] Needed: Better ways to use world's technical literature. «SAE Journal», 1958, v. 66, № 9, стр. 60; President approves program for dissemination of scientific information. «Science Information News», 1959, v. 1, № 1, стр. 2.

нических отчетов, причем их число увеличивается на 30—40% в год [36]. Кроме того, ежегодно в мире издается до 75 тыс. книг по науке и технике [37] и публикуется более 300 тыс. описаний к авторским свидетельствам и патентам. По ориентировочным данным, общий объем мировой научно-технической литературы в настоящее время удваивается каждые 8,5 лет. В течение последних 100 лет фонды научно-технической литературы в Библиотеке конгресса США удваивались примерно каждые 20 лет и в настоящее время содержат свыше 1,5 млн. томов книг и журналов [38].

В последнее время начали действовать еще два важных фактора, существенно затрудняющих для исследователя задачу самостоятельного ознакомления с интересующей его научно-технической литературой. Один из них заключается в том, что ликвидация колониальной системы империализма вывела на путь самостоятельного политического, экономического и культурного развития десятки новых стран. В этих странах начали создаваться национальные научные кадры и шире развертываться научно-исследовательские и опытно-конструкторские работы. Результатом этого процесса является непрерывный рост в мировой научно-технической литературе удельного веса важных публикаций, выходящих на языках этих стран. Еще до второй мировой войны важнейшие научно-технические публикации выходили на нескольких основных языках, и большинство ученых и инженеров могло самостоятельно знакомиться с ними. Теперь положение коренным образом изменилось. По произведенным в США подсчетам, американским ученым в языковом отношении доступно для самостоятельного ознакомления не менее 1/3 всей мировой научно-технической литературы. Но таких полиглотов среди ученых США насчитывается менее 5%. Если судить по американскому реферативному журналу «Chemical Abstracts», то более 30% всех материалов по химии выходит на русском, польском, венгерском, чешском, болгарском, китайском, итальянском и японском языках. Научную литературу на этих языках могут читать менее 2% американских ученых [39].

Другим фактором является прогрессирующее стирание привычных границ между традиционными областями знания, углубляющееся взаимопроникновение отдельных наук. Теперь исследователь не может ограничиться ознакомлением с научно-техни-

---

[36] Там же, стр. 2.

[37] C. P. Bourne. Указ. соч., стр. 1.

[38] Documentation, indexing, and retrieval of scientific information. Report, Committee on Government Operations. U. S. Senate, 86th Congress, 2nd Session. Washington, U. S. Government Printing Office, 1960, стр. 116.

[39] Chemical literature expands. «Chemical and Engineering News», 1966, v. 44, № 23, стр. 86; D. B. Baker. Growth of chemical literature. Past, present, and future. «Chemical and Engineering News», 1961, v. 39, № 29, стр. 79—80.

ческой литературой лишь по одной отрасли знания. Чтобы вести плодотворную научную работу, он должен знать о результатах исследований, выполненных в других областях науки и техники.

В связи с этим нельзя не отметить, что публикуется значительно больше научно-технической литературы, чем это необходимо для достаточно полного изложения результатов выполненных научно-исследовательских и опытно-конструкторских работ. Одной из причин такой ненужной избыточности научно-технической литературы является уже устаревшая традиция, по которой число опубликованных исследователем работ принято считать формальным показателем уровня его научной квалификации. Указывая на это, Норберт Винер писал: «При колоссальном объеме публикаций в последние годы молодой ученый боится подорвать свою карьеру из-за того, что у него нет достаточно ходких результатов, чтобы заполнять ими журналы или чтобы удовлетворить требования своих руководителей, стремящихся, в частности, иметь продукцию от своих институтов» [40]. Многие ученые, обеспокоенные быстрым количественным ростом научно-технической литературы, в последнее время все чаще предлагают различные меры, направленные на то, чтобы пресечь повторение публикаций одних и тех же материалов в разных изданиях.

Быстрый количественный рост документальной научно-технической информации ведет к тому, что исследователю приходится затрачивать все бо́льшую часть своего рабочего времени не на исследования, а на ознакомление с соответствующей научно-технической литературой и поиски нужной ему информации. В 1958 г. Технологический институт Кейса совместно с Американским химическим обществом и Национальным научным фондом США провел статистическое обследование баланса рабочего времени химиков. Этим обследованием было охвачено около 1500 химиков, работающих в 45 фирмах и 5 университетах. В зависимости от характера деятельности они были разделены на 50 групп. В ходе обследования было сделано примерно 25 тыс. наблюдений. Общие результаты этого обследования показаны на рис. 10 [41].

Условно было выделено 8 основных видов деятельности, которой химики заняты в рабочее время. По каждому из этих видов, указанных на рис. 10, отмечены минимальные и максимальные затраты рабочего времени (в процентах), а также подсчитаны их средние значения. Естественно, что лишь сумма этих средних значений составляет 100%. Из сопоставления этих средних значе-

---

[40] Н. Винер. Наука и общество. «Вопросы философии», 1961, № 7, стр. 122.

[41] M. H. Halbert, R. L. Ackoff. An operations research study of dissemination of scientific information. В кн.: «Preprints of papers' for the International Conference on Scientific Information». Washington, November 16—21, 1958. Washington, National Academy of Sciences—National Research Council, 1958, Area 1, стр. 87—120.

ний затрат рабочего времени на разные виды деятельности видно, что в настоящее время американский химик расходует на все виды информационной деятельности уже около 44% своего рабочего времени, тогда как на экспериментальную работу — всего 36%. Под информационной деятельностью в данном случае понимается чтение специальной научно-технической литературы и поиск в ней нужных сведений, устные сообщения и подготовка соответствующих письменных документов — отчетов, статей, обсуждение различных научных и служебных вопросов и т. д. Примерно такое же положение складывается и в других отраслях науки и техники.

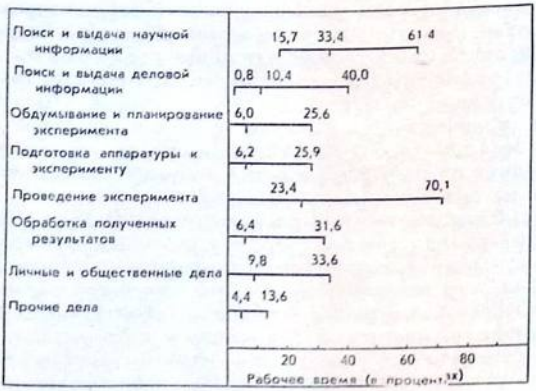



Рис. 10. Распределение рабочего времени американских химиков

Однако дальнейшее увеличение доли рабочего времени, которую современный исследователь вынужден расходовать на научно-информационную деятельность, — не единственный недостаток современной системы распространения научных знаний. Ныне исследователь или даже группа исследователей уже не в состоянии следить за всей научно-технической литературой в интересующей их области. Это неизбежно ведет к повторению уже выполненных и описанных в литературе научно-исследовательских и опытно-конструкторских работ.

Вместе с тем научные исследования с каждым годом приобретают все более сложный характер и поэтому требуют для своего проведения больших творческих усилий многочисленных коллективов ученых и инженеров, времени и материальных средств. В таких условиях неоправданное дублирование исследований

и разработок оказывает отрицательное влияние на общие темпы развития науки и техники и дорого обходится человечеству.

На один из экспериментов по засеву облаков в США было израсходовано 256 тыс. долларов. Результаты этого эксперимента были опубликованы. Однако на его ненужное повторение было израсходовано еще 3 млн. долларов, пока отчет о ранее проведенном эксперименте не был случайно обнаружен в библиотеке [42]. Национальное управление по аэронавтике и исследованию космического пространства и Министерство ВВС США израсходовали год времени и 18 млн. долларов на параллельную разработку аналогичных ракетных систем — «Атлас-Вега» и «Атлас-Агена В» [43].

В письме к редактору американского журнала «Science», опубликованному 11 мая 1962 г., было обращено внимание на тот факт, что третий пояс радиации, расположенный на высоте 64 тыс. *км* от Земли, был открыт в 1960 г. советским ученым К. И. Грингаузом. Информация об этом была опубликована в февральском и апрельском номерах «Докладов Академии наук СССР» за 1960 г. Однако большинство американских ученых, работающих в той же области, по-видимому, не знали об этих публикациях, так как в 1961 г. в США было объявлено, что американский искусственный спутник Земли «Эксплорер-12» обнаружил на такой же высоте третий пояс радиации, причем никаких упоминаний об открытии советского ученого сделано не было [44].

За последние годы в архивных учреждениях только нашей страны было обнаружено более 40 тыс. документов о месторождениях полезных ископаемых, которые были потом забыты [45].

Еще один пример. Учреждения, работавшие над проектированием Северо-Крымского канала, даже не поинтересовались весьма ценными материалами экспедиции для изыскания средств к обводнению и облесению Крыма, работавшей еще в 1865—1866 гг. Эти материалы хранятся в Центральном государственном архиве в Ленинграде [46].

В 50-х годах у нас в стране выходило много литературы о кукурузе, причем авторы часто описывали ее как малоизученную

---

[42] E. Herbert. Finding what's known. «International Science and Technology», 1962, № 1, стр. 14.

[43] P. Fozzy. Research coordination. «Bull. of the Atomic Scientists», 1962, v. 17, № 3, стр. 31.

[44] F. Coleman. Keeping up with Soviet research. «Science», 1962, v. 136, № 3515, стр. 594—596; К. И. Грингауз и др. Изучение межпланетного ионизированного газа, энергичных электронов и корпускулярного излучения Солнца при помощи трех электродных ловушек заряженных частиц на второй советской космической ракете. «Докл. АН СССР», 1960, т. 131, № 6, стр. 1301—1304.

[45] А. Горбовский. Старые загадки истории и новые гипотезы. «Наука и жизнь», 1963, № 6, стр. 102.

[46] Г. Севунц. Подвиг эпохи (Изобретатели велосипедов). «Москва», 1960, № 4, стр. 179.

и незнакомую нам культуру. Но если заглянуть в литературу 80-х годов XIX в., то можно обнаружить удивительные вещи. Оказывается, что еще в конце прошлого столетия культура кукурузы была в России хорошо освоена и на европейском рынке русское кукурузное зерно успешно конкурировало с американским.

Таких примеров можно привести множество. По ориентировочным данным, в настоящее время на неоправданное дублирование в США расходуется не менее 10% всех средств, ассигнуемых в стране на научно-исследовательские и опытно-конструкторские работы [47]. Такое дублирование имеет место, несмотря на то, что в США ежегодно расходуется свыше 1 млрд. долларов на документальный информационный поиск [48]. По другим данным, американские ученые тратят на неоправданное дублирование от 30 до 80% своего рабочего времени [49].

Приведенные цифры свидетельствуют о том, что традиционные библиотечно-библиографические методы и средства, которые в настоящее время все еще доминируют в научно-информационной деятельности, более не являются достаточно эффективными и что стоимость информационного поиска, основанного на применении этих методов и средств, быстро растет. В связи с этим следует отметить, что иногда стоимость такого информационного поиска бывает настолько высокой, что делает бессмысленным его проведение. Поэтому не случайно, что в настоящее время некоторые американские фирмы считают, что если научно-исследовательская или опытно-конструкторская работа стоит менее 10 тыс. долларов, то им выгодней повторить эту работу, чем выяснить, была ли она выполнена где-либо ранее и опубликованы ли ее результаты в научно-технической литературе [50]. В 1966 г. в нашей стране как уже отмечено, насчитывалось 711,6 тыс. научных работников [51]. Нетрудно подсчитать, что если лишь наполовину разгрузить их от информационной работы, на которую сейчас уходит 1/3 рабочего времени, то это было бы эквивалентно увеличению общего числа наших научных работников на 118,6 тыс. человек. Разумеется, в данном случае мы имеем в виду лишь техническую

---

[47] F. Bello. How to cope with information. «Fortune», 1960, v. 72, № 3, стр. 163—164.

[48] Information retrieval $ 200 million in 1968. «Electronics», 1961, v. 34, № 46, стр. 12.

[49] I. Hirsch, W. Milwitt. Increasing the productivity of scientists. «Harvard Business Review», 1958, v. 36, № 2, стр. 69.

[50] P. C. Constant Jr., P. C. Patton. A manual information retrieval system. «IRE Transactions on Engineering Management», 1962, v. 9, стр. 87. — Вице-президент одной из сталелитейных корпораций США, вследующий научно-исследовательской работой, называл в таком же контексте цифру в 100 тыс. долларов. См.: F. T. Sisco. What's literature? «Metal Progress», 1957, v. 72, № 4, стр. 124.

[51] СССР в цифрах в 1966 г. Краткий статистический сборник. М., «Статистика», 1967, стр. 160.

сторону работы, а не освобождение исследователя от чтения научной литературы, которое всегда будет важной составной частью исследования. Приведенные цифры будут еще более внушительными, если учесть, что информационной работой в большей или меньшей степени вынуждены заниматься и многие из наших инженеров. В 1966 г. число дипломированных инженеров в СССР составляло 1789 тыс. человек [52]. Все это ясно свидетельствует о том, что в настоящее время частичное освобождение исследователя от необходимости расходовать рабочее время на поиски научно-технической литературы является огромным резервом повышения производительности его научного труда.

Непрерывное увеличение количества научно-технической литературы, а также расширение и усложнение информационных запросов исследователей создали такое положение, когда достаточно быстрая и исчерпывающая информация может быть обеспечена лишь при условии разработки и широкого применения принципиально новых методов и средств, специально предназначенных для переработки документальной информации. Выделение научно-информационной деятельности в самостоятельную часть научного процесса ведет к формированию новой научной дисциплины и новой категории специалистов, для которых основным объектом научного исследования являются закономерности информационного процесса, научные документы как материальные носители закрепленной информации и специальные средства, используемые в информационном процессе.

Однако было бы упрощением считать, что проблема научной информации возникла лишь в результате быстрого количественного роста научно-технической литературы. В одной из своих статей И. Бар-Хиллел справедливо отмечал, что при оценке проблем, которые создает количественный рост научно-технической литературы, нельзя забывать об углубляющейся дифференциации наук. Вследствие постоянного сужения области, в которой работает каждый ученый или инженер, количество интересующих его научно-технических публикаций растет значительно медленней, чем общее количество научно-технической литературы [53]. Но из признания такого, в известной степени компенсирующего действия специализации вовсе не следует, что проблемы научной информации вообще не существует. Это лишь означает, что существуют и другие важные факторы, которые делают эту проблему научной информации одной из наиболее актуальных проблем нашего времени.

К числу таких факторов следует прежде всего отнести две особенности, характерные для современного этапа развития науки

---

[52] СССР в цифрах в 1966 г. Краткий статистический сборник. М., «Статистика», 1967, стр. 51.

[53] J. Bar-Hillel. Is information retrieval approaching a crisis? «American Documentation», 1963, v. 14, № 2, стр. 95—98.

и техники. Первая из этих особенностей заключается в том, что в настоящее время научно-исследовательская и опытно-конструкторская работа стала массовой специальностью. Этого не было еще 30—40 лет назад. Вовлечение в сферу деятельности, непосредственно связанную с развитием науки и техники, все большей части общества породило общественную потребность в организации научно-информационного обслуживания ученых и инженеров. Более того, научно-информационное обеспечение само приобрело характер задачи массового обслуживания, для эффективного решения которой необходимо применение средств механизации и автоматизации.

Вторая особенность современного этапа развития науки и техники — это резкое сокращение сроков между научными открытиями и их внедрением в повседневную практику. Именно это обстоятельство создает впечатление исключительности характера современной научно-технической революции. В своей статье «Естественные науки и их значение для развития мировоззрения и технического прогресса» академик М. В. Келдыш писал: «Когда говорят, что именно сейчас наступил век науки и техники, то это не совсем правильно. Ведь во все времена материально-технический прогресс был в значительной мере основан на развитии науки. Корни промышленной революции также лежали в науке. Прошлое столетие характеризовалось очень большим количеством крупнейших научных достижений. Неверно думать, что раньше было мало открытий в области естествознания, повлекших за собой большие последствия в области материального производства, что только сейчас мы вступили в век непрерывных открытий.

Однако есть одна черта, весьма характерная для современного развития науки и техники: это быстрота практического использования научных открытий. Можно привести ряд примеров. С момента открытия электрического тока (Гальвани) до создания первой электростанции прошло около века. Потребовалось почти сто лет, чтобы освоить это замечательное открытие, имевшее громадные перспективы. Можно также отметить, что со времени выяснения роли минеральных удобрений в питании растений (середина прошлого века) и до их интенсивного использования прошло в общем лет семьдесят. Но они стали широко применяться только после второй мировой войны. Иначе обстояло дело с открытием деления ядра урана. С момента этого открытия до создания ядерного реактора прошло всего три года, а до создания первой атомной электростанции — 15 лет.

Именно потому, что наше время характеризуется чрезвычайно быстрым использованием достижений науки, все более важное значение приобретает хорошая организация научных исследований и использование их результатов в производстве. Сегодня в реализации производственного процесса может оказаться впереди не та страна, которая первой сделает новое научное открытие,

а та, которая сможет лучше организовать его быстрейшее использование в практике» [54]. Быстрое сокращение сроков исследований и разработок, а также внедрения открытий и изобретений в практику графически показано на ряде других примеров на рис. 11.

Такое ускорение темпов исследований и разработок требует соответствующего увеличения скорости действия информационных систем. Очевидно, что эта задача может быть решена лишь путем широкого применения в информационной практике принципиально новых методов работы, а также использования соответствующих средств механизации и автоматизации.

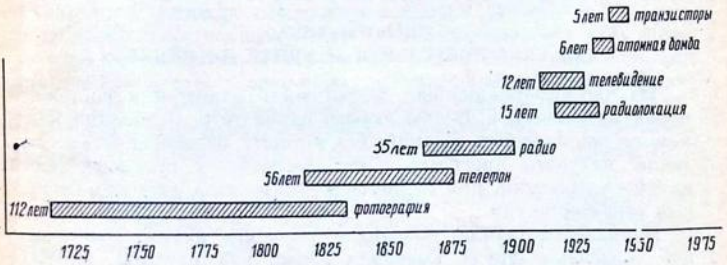



Рис. 11. Сокращение сроков между научными открытиями и их внедрением в практику

Практический опыт подтверждает, что производительность труда исследователей значительно возрастает, когда подбор и предварительный анализ нужной им информации производятся научными работниками, специально подготовленными для выполнения этой разновидности научного труда. Поэтому подготовка научных работников такого профиля приобретает исключительно большое значение для повышения темпов развития современной науки. На это указывалось, например, в докладе «Наука, правительство и информация», который в январе 1963 г. был представлен президенту США Дж. Ф. Кеннеди его Консультативным комитетом по научным вопросам. В докладе особо подчеркивалась необходимость того, чтобы специалисты по научной информации получили в науке такой же статут, какой имеют в ней исследователи. Это мотивировалось тем, что «информационный процесс является интегральной частью научно исследовательских и опытно-

[54] М. В. Келдыш. Естественные науки и их значение для развития мировоззрения и технического прогресса. «Коммунист», 1966, № 17, стр. 33. Данные для рис. 11 заимствованы из издания: The future is our assignement. A glimpse behind the scenes at the General Motors Research Laboratories. Detroit, 1959, стр. 5.

конструкторских работ» и что «. . . в будущем научно-информационная деятельность может занять в науке примерно такое же место, какое занимает сегодня в физике теоретическая физика» [55].

Коммунистическая партия и Советское правительство всегда придавали научной информации большое значение.

В Программе КПСС, принятой на XXII съезде партии, говорится: «Партия будет всемерно содействовать . . . образцовой постановке научно-технической информации, всей системы изучения и распространения отечественного и зарубежного передового опыта» [56].

## ИНФОРМАТИКА КАК САМОСТОЯТЕЛЬНАЯ НАУЧНАЯ ДИСЦИПЛИНА

Научно-информационная деятельность является в настоящее время органической частью любого научного исследования и одним из непременных условий, без которых невозможны высокие темпы научного прогресса. Она обеспечивает преемственность научного развития, при ее помощи реализуется интернациональный характер науки.

Накопление знаний, вызванное развитием науки, общественное разделение труда выделили новый вид научной деятельности — научно-информационную деятельность, которая складывается в условиях сложных процессов, протекающих в науке, — глубокой дифференциации и синтеза знаний, внедрения в широких масштабах экспериментальных приемов изучения, использования мощных технических средств и растущего проникновения математических методов в различные области науки. При разработке сложных научно-технических проблем возникает необходимость объединения большого числа ученых самых различных специальностей, что приводит к массовости научных исследований.

Общественное разделение труда в науке происходит не только по линии формирования новых отраслей науки, но и по функциональному признаку внутри каждой науки. Таким образом, выделились, например, теоретическое и экспериментальное направления. Одновременно происходит выделение и третьего самостоятельного вида научного процесса — научно-информационной деятельности.

При современных темпах развития науки и внедрения ее достижений в практику достаточно быстрая и исчерпывающая информа-

---

[55] Science, government, and information. A report of the President's Science Advisory Committee. Washington, U. S. Government Printing Office, 1963, стр. 2, 14—15, 28—29.

[56] Программа КПСС. М., Госполитиздат, 1962, стр. 159—160.

ция может быть обеспечена лишь при условии разработки и широкого применения принципиально новых методов и средств, специально предназначенных для переработки документальной информации. Традиционные библиографические методы обработки, хранения и поиска документальной информации в современных условиях становятся малоэффективными. Несмотря на достаточно высокий научный уровень современной библиографии и библиотековедения и большой опыт работы в этом направлении, научно-информационная деятельность, используя в известной мере этот опыт, должна разрабатывать свои собственные методы и средства.

Современная научно-информационная деятельность должна обеспечивать решение следующих основных задач:

1. Предельно полный сбор, аналитико-синтетическую переработку документальной информации с целью быстрого оповещения исследователей о всех новых достижениях науки и техники.

2. Долговременное хранение документальной информации в информационно-поисковых системах, позволяющих осуществлять быстрый, исчерпывающий и многоаспектный поиск необходимых сведений.

3. Переработку информации в информационно-логических системах с целью получения новой информации, например, нахождение на основе записанных в машинной памяти «химических аналогий» вероятных путей синтеза еще не полученных химических соединений.

Первые две задачи в какой-то степени могут быть решены и частично уже решаются традиционными методами и средствами, бóльшая часть которых носит эмпирический и полуэмпирический характер. Однако и чисто практические задачи могут быть успешно решены в настоящее время лишь на основе разработки специальной теории. Теоретическое (а не только техническое) обоснование должны получить практика и перспективы использования средств механизации и автоматизации.

Необходимость применения принципиально новых методов и средств отчетливо видна при рассмотрении третьей задачи информационной деятельности — задачи логической переработки научной информации. Традиционными средствами и методами эта задача вообще не может быть решена. Между тем ее решение имеет исключительно большое значение, так как позволяет получать новую информацию путем логической переработки уже имеющейся информации.

Информационные задачи, возникающие перед исследователем, могут носить самый разнообразный характер. Некоторые из этих задач отличаются достаточно высоким уровнем сложности. Примером может служить поиск химических соединений с нужными физико-химическими свойствами по соответствующим фрагментам их структурных формул.

Допустим, что нам нужно отыскать химические соединения, обладающие некоторыми заданными физико-химическими свойствами. Весьма вероятно, что соединение или соединения с такими свойствами уже имеются в числе известных и описанных в литературе. Эту задачу можно было бы решить путем изучения структурных формул всех известных соединений данного класса и отбора тех из них, которые содержат характерные фрагменты, придающие соединениям интересующие нас свойства. Однако следует напомнить, что в серии «Химия» реферативного журнала ВИНИТИ ежегодно помещается информация более чем о 180 тыс. химических соединений, из которых около половины предположительно являются новыми соединениями. Всего в настоящее время учтено около 2,5 млн. неорганических и органических соединений. Ясно, что если химик будет пользоваться лишь такими традиционными средствами, как справочники, картотеки, реферативные журналы с указателями и т. п., то он не сможет в приемлемые сроки решить эту задачу.

Решение подобных задач можно поручить специализированной информационной системе кибернетического типа, которая способна производить нужные операции сопоставления за ничтожные доли секунды. Предварительно в запоминающее устройство такой информационной системы должны быть введены линейные шифры структурных формул всех известных химических соединений или соответствующего класса соединений. Она сможет решать и другие сложные информационные задачи, связанные с логической переработкой имеющейся информации. Информационная система, например, может определять физические свойства вещества при заданных условиях, если в литературе описаны физические свойства данного вещества при других условиях; отыскивать возможные пути синтеза определенных соединений из некоторого набора исходных продуктов на основе аналогий, введенных в запоминающее устройство информационной системы.

Практика последних лет, большое число научно-исследовательских работ, ведущихся у нас в стране и за рубежом по проблемам научно-информационной деятельности, все большее число специалистов, вовлекаемых в эту сферу, вызвали широкое развитие теоретической мысли в этой области. Начала формироваться новая научная дисциплина, изучающая принципы информационной деятельности и на этой основе разрабатывающая эффективные методы и средства сбора, переработки, хранения и распространения документальной информации. Научно-информационная деятельность, как интегральная часть научных исследований, может быть эффективной лишь тогда, когда она выделена в самостоятельную область и имеет свою теоретическую основу.

Известно, что ни одна научная дисциплина или отрасль науки не возникает на пустом месте. Движущей силой развития науки и критерием истинности ее положений является общественная

практика людей. К настоящему времени в научно-информационной деятельности уже накоплено большое количество фактов.

Ф. Энгельс в «Диалектике природы» подчеркивал, что «в любой научной области — как в области природы, так и в области истории — надо исходить из данных нам *фактов* . . . нельзя конструировать связей и вносить их в факты, а надо извлекать их из фактов и, найдя, доказывать их, насколько это возможно, опытным путем» [57]. Однако факты, взятые сами по себе, представляют лишь груду сырого материала и не составляют науки. С помощью теоретического мышления, научной абстракции человек, отвлекаясь от случайных связей явлений, проникает в глубинные связи, открывает объективные законы, которые лежат в основе процессов природы, общественной жизни и мышления. Вот почему в научно-информационной деятельности нельзя ограничиваться лишь накоплением фактов, необходимо изучать их взаимные связи, вскрывать внутренние закономерности научно-информационных процессов и на этой основе строить соответствующую научную теорию.

Появление новой научной дисциплины, именуемой «документацией», «документалистикой», «научной информацией», «теорией научной информации», «информационной наукой» (information science) и т. д., в настоящее время является общепризнанным фактом, оспаривать который не будет ни один серьезный исследователь. Эта новая научная дисциплина возникла в ответ на все увеличивающуюся общественную потребность в обеспечении эффективных методов и средств для сбора, обработки, хранения, поиска и распространения научной информации. Подчеркивая данное обстоятельство, известный американский специалист в этой области Р. Тейлор писал:

«Записанные (recorded) знания, необходимые для проведения научно-исследовательских и опытно-конструкторских работ в технически развитом обществе, слишком рассеяны и распылены, чтобы с ними могли справиться обычные библиотечные системы. Новые научные дисциплины, возникшие за последние 20 лет, не только предъявили непредвиденные требования к библиотекам, но и породили совершенно новый комплекс идей о процессе передачи информации.

С появлением этих проблем возникла и специальная наука об информации и документации. Как отрасль науки она исследует структуру, свойства и процесс передачи информации, пользуясь для этого методами таких наук, как психология, логика, нейрофизиология, лингвистика и математика. Как отрасль техники она имеет дело со средствами обработки информации, производимой с целью обеспечения ее оптимальной доступности и использо-

---

[57] Ф. Энгельс. Диалектика природы. М., Госполитиздат, 1955, стр. 26.

вания, и применяет методы вычислительной техники, библиотековедения, системотехники и наук об управлении» [58].

Первоначально для обозначения этой новой научной дисциплины был использован термин «документация», ведущий свое происхождение от слова «документ». Одно из первоначальных значений термина «документ» было чисто юридическим — письменное подтверждение прав и состояний в предписанной законом форме. Другое, более позднее и широкое значение этого термина — материальный объект, содержащий закрепленную информацию, специально предназначенный для ее передачи в пространстве и времени и используемый в общественной практике.

Совокупность документов обычно называется документацией. В этом смысле принято говорить о технической документации, торговой документации и т. д. В качестве термина, обозначающего специальную деятельность по сбору, обработке, хранению, поиску и распространению документов, отличную от библиотечного дела и библиографии, термин «документация» был употреблен известным бельгийским ученым и общественным деятелем Полем Отле в 1905 г. в докладе на Международной конференции по экономике [59]. В 1920 г. при Патентном управлении Нидерландов был создан Голландский институт документации [60]. С тех пор этот термин в указанном выше значении получил широкое распространение. В 1934 г. он вошел в название Международного института документации, в который был преобразован Международный библиографический институт, а в 1937 г. — в название организованной на его основе Международной федерации документации (МФД) [61].

В долгосрочной программе МФД документация определяется «как сбор и хранение, классификация и подбор, распространение и использование всех видов информации» [62]. По мнению многих теоретиков документации, одним из основных ее признаков является способность быстро отражать новые сведения, объединять разрозненные материалы, обобщать информацию, содержащуюся в опубликованных и неопубликованных документах, отвечать на узкоспециальные и многоаспектные запросы.

---

[58] R. Taylor. Glossary of terms frequently used in scientific documentation. Seattle, 1962, стр. VI.

[59] P. Otlet. L'organisation rationnelle de l'information et de la documentation en matiere economique. Bruxelles, 1905.

[60] Позже этот институт был переименован в Нидерландский институт документации и архивного дела.

[61] История появления термина «документация» в названии МФД изложена в статьях: F. Donker Duyvis. Die Entstehung des Wortes «Dokumentation» in Namen der FID. «Revue de la documentation», 1959, v. 26, № 1, стр. 15—16; C. Björkbom. The history of the word documentation within the FID. Там же, № 3, стр. 68—69.

[62] Outline of a long-term policy of the International Federation for Documentation. FID publ. 325. The Hague, FID, 1960, стр. 9.

В течение довольно долгого времени велись поиски наиболее точного определения документации. При этом обращалось особое внимание на специфические особенности, отличающие документацию от библиотечного дела и библиографии.

«Документация — это искусство сбора, классификации и быстрого отыскания записей о всех видах интеллектуальной деятельности. Это процесс, который позволяет . . . предоставлять исследователю существующую литературу по предмету его исследования с тем, чтобы он имел возможность знать о всех достижениях в его области и таким образом не растрачивать напрасно своих творческих сил (genius) на повторение уже выполненных работ» (С. Брэдфорд, 1948) [63].

«Документация — это такая часть библиографической организации, которая занята опосредованной передачей первичных материалов внутри и среди групп специалистов с тем, чтобы они получали в возможно более действенной форме сведения, необходимые им для эффективного выполнения их работы» (Дж. Шира, 1951) [64].

Документация — это «обозначение всего комплекса операций, осуществляемых в процессе передачи специализированной информации; она включает все виды деятельности, составляющие библиотечное дело плюс предшествующие им операции по подготовке и воспроизведению материалов, а также последующую деятельность по их распределению» (М. Таубе, 1953) [65].

«Документация — это совокупность (group) методов, необходимых для упорядоченного представления, систематизации и передачи записанных специализированных знаний с целью обеспечения максимальной доступности и использования содержащейся в этих записях информации» (Дж. Мак и Р. Тейлор, 1956) [66].

«Документация — это сбор, кодирование и распространение записанных знаний, рассматриваемые в целом как интегральная процедура, использующая семантику, психологические и механические средства, а также методы репродуцирования, включая микрокопирование, с целью обеспечения максимальной доступности и использования документальной информации» (Словарь Вебстера, 1961) [67].

---

[63] S. C. Bradford. Documentation. 2nd. ed. London, Lookwood, 1953, стр. 49.

[64] J. Shera. Documentation: its scope and limitations. «Library Quarterly», 1951, v. 21, № 1, стр. 19.

[65] M. Taube. Theory and practice of documentation. В кн.: M. Taube et al. Studies in coordinate indexing. V. 1. Washington, Documentation, 1953, стр. 7.

[66] J. D. Mack, R. S. Taylor. A system of documentation terminology. В кн.: J. H. Shera, A. Kent, J. W. Perry (eds). Documentation in action. New York, Reinhold, 1956, стр. 20.

[67] Webster's Third new international dictionary. Cleveland—New York, 1961.

Документация «занимается анализом информационных систем и разработкой методов для новых систем» (Дж. Бейкер и Р. Хейс, 1963) [68].

Ю. Джексон считал наиболее удачным определение, официально принятое в 1951 г. исполнительным бюро Ассоциации специальных библиотек США: «Документация — это искусство а) создания документов, б) распространения документов и в) использования документов» [69]. В этом определении под документом подразумевается любая записанная информация, независимо от формы ее записи и способа использования. Это могут быть отчеты, фильмы, фотографии, географические карты, перфокарты, магнитные ленты, осциллограммы и т. п.

В 1955 г. Американский институт документации провел конкурс на лучшее определение термина «документация». Наилучшими были признаны следующие определения:

«Наука об упорядоченном представлении и сохранении записанных научных знаний (records of knowledge), призванная раскрывать эти знания в форме, удобной для их быстрого использования и установления связей между ними» (Г. Дайсон).

«Процедура, благодаря которой накопленный запас знаний можно использовать для получения новых знаний» (А. Сейделл).

«Искусство, облегчающее использование записанных специализированных знаний посредством их представления, репродуцирования, опубликования, распространения, сбора, хранения, смыслового анализа, организации и поиска» (Э. Браунсон) [70].

Наряду с приведенными определениями документации, каждое из которых вносит какой-либо нюанс в трактовку содержания этой научной дисциплины, имеется много других определений, в той или иной мере совпадающих с цитированными здесь. Они принадлежат таким крупным специалистам, как М. Верхуф [71], Э. Пич [72], Т. Лошес [73], В. Янсен [74], Ф. Вагнер [75] и др.

---

[68] J. Becker, R. M. Hayes. Information storage and retrieval. Tools, elements, theories. New York, Wiley, 1963, стр. 44.

[69] E. B. Jackson. Inside documentation. «Special libraries», 1954, v. 45, № 4, стр. 151.

[70] Definitions of documentation. «American Documentation», 1955, v. 6, № 4, стр. 254.

[71] М. Верхуф. Библиотечное дело и документация. «Бюлл. ЮНЕСКО для библиотек», 1960, т. 14, № 5, стр. 213—216.

[72] E. Pietsch. Grundfragen der Documentation. Dortmund, 1954. Sonderdruck aus Heft 14 der «Schriftenreihe der Arbeitsgemeinschaft für Rationalisierung des Landes Nordrhein—Westfalen».

[73] T. P. Loosjes. Documentation van weterschappelijk literatuur. Amsterdam, 1957.

[74] W. H. J. Jansen. Het begrip documentatie. «Bibliotheek-leven», v. 45, № 10.

[75] F. S. Wagner. A dictionary of documentation terms. «American Documentation», 1960, v. 11, № 2, стр. 108.

В последнее время в зарубежной литературе иногда употребляется термин «научная документация».

Например, Национальный научный фонд США с 1957 г. выпускает полугодовые обзорно-реферативные сборники под заглавием «Текущие исследования и разработки в области научной документации» [76]. Фундаментальная монография румынских ученых А. Аврамеску и В. Кындя вышла под заглавием «Введение в научную документацию» [77]. Термин «научная документация» употреблен также в названии цитированного нами словаря Р. Тейлора [78]. Однако этот термин в принципе мало чем отличается от термина «документация».

Параллельно с термином «документация» в англо-американской литературе с конца 40-х годов начинает употребляться термин «научная информация» для обозначения того же вида деятельности и той же научной дисциплины. В 1948 г. он вошел в название Конференции Королевского общества по научной информации, проходившей в Лондоне, а в 1958 г. — в название Конференции по научной информации, проходившей в Вашингтоне.

В 1959 г. в ГДР возникла дискуссия относительно употребления этих двух терминов и взаимосвязи между ними. В ходе этой дискуссии Й. Коблитц высказал мнение, что документация включает лишь переработку содержания документов, являющуюся предпосылкой для дальнейшей информации. По его мнению, документация заключается в установлении пригодности документов, их переработке и классификации, а информация — в хранении, поиске и распространении документов [79]. С ним не согласился И. Ползович (ВНР), который доказывал, что научная информация является одной из функций документации, ее составной частью. Он писал, что «в ходе развития документации как отрасли науки и практической деятельности выделились определенные процессы, обозначаемые термином «информация»; однако эти процессы не выходят за рамки документации и сохраняют с ней органическую связь» [80].

Д. Исакович из Югославского центра технической и научной документации недавно проанализировал ряд высказываний

---

[76] Current research and development in scientific documentation. № 1—14. Washington, U. S. Government printing office, 1957—1966.

[77] A. Avramescu şi V. Cândeă. Introducere in documentarea ştiintifică. Bucureşti, Acad. Republicii Populare Romîne, 1960.

[78] R. Taylor. Glossary of terms frequently used in scientific documentation. Seattle, 1962.

[79] J. Koblitz. Dokumentation und Information. Eine terminologische Untersuchung. «Dokumentation», 1959, v. 6, № 1.

[80] I. Polzovicz. Einige Betrachtungen zur Begriffsbestimmung der Termini Dokumentation und Information. «Dokumentation», 1959, v. 6, стр. 171.

относительно названий новой научной дисциплины [81]. Изложив позиции авторов ряда работ, опубликованных в сборнике «Научно-техническая информация» (А. И. Михайлов и В. А. Полушкин), «Бюллетене ЮНЕСКО для библиотек» (М. Верхуф и Г. Яхода), в журнале «Dokumentation» (Й. Коблитц), Д. Исакович пришел к следующим выводам.

Складывающаяся в настоящее время научная дисциплина, задачи и методы которой сформулированы в уставе МФД, имеет два разных названия — «теория научной информации» и «документация», — которые можно считать синонимами. «Мне представляется, — пишет Д. Исакович, — что предлагаемый термин «теория научной информации» не является полностью адекватным. По моему мнению, с точки зрения содержания этот термин сужает широкую область, о которой идет речь. С другой стороны, термин «документация» не раскрывает существа той научной дисциплины, которую он призван обозначать» [82]. В заключение автор этой статьи, опубликованной в официальном органе МФД, предлагает ввести для обозначения рассматриваемой научной дисциплины термин «документология».

Возможно, не было бы необходимости уделять здесь столько внимания термину «документация», поскольку он не нашел применения в нашей стране. В ходе развития информационного дела в СССР не возникло практической или теоретической необходимости искусственно разделять единую область деятельности и научную дисциплину на два самостоятельных или даже взаимосвязанных комплекса. Из двух параллельно существующих терминов «документация» и «информация» был выбран лишь один. То, что выбор пал на второй из них, можно объяснить справедливым стремлением подчеркнуть, что проблемы научно-информационной деятельности выходят за рамки документа, т. е. не обязательно связаны с документом.

Однако в советской специальной литературе, наряду с термином «научная информация», получил также распространение и термин «документалистика». Он был впервые предложен В. А. Успенским в 1958 г. на заседаниях семинара по теоретическим основам научной информации в Лаборатории электромоделирования ВИНИТИ АН СССР [83] для обозначения того раздела дисциплины, в котором изучается документальная информация. Следует отметить, что В. А. Успенский в статье «К проблематике теории научной информации», написанной им в 1963 г. совместно с Ю. А. Шрей-

---

[81] D. Isaković. La documentologie. «Revue internationale de la documentation», 1965, v. 32, № 4, стр. 152—153.

[82] Там же, стр. 153.

[83] Н. И. Стяжкин и Н. Д. Кравченко. К проблематике документалистики. «Докл. на конференции по обработке информации, машинному переводу и автоматическому чтению текста». Вып. 4. М., 1961, стр. 2.

дером, называет соответствующую дисциплину теорией научной информации [84].

Тем не менее некоторые специалисты продолжают до настоящего времени использовать термин «документалистика» для обозначения новой научной дисциплины. Например, в 1966 г. издательство «Наука» выпустило сборник статей, который составители озаглавили «Кибернетика и документалистика» [85].

Термин «документалистика» неудовлетворителен по двум причинам. Во-первых, этот термин, как и термин «документация», излишне акцентирует внимание на документе, тогда как предметом исследования рассматриваемой дисциплины являются структура и свойства научной информации. Во-вторых, он образован на основе термина «документалист», используемого для обозначения специалиста в области документации, и замыкает ряд терминов, возникавших в следующей последовательности: «документ» — «документация» — «документалист» — «документалистика». Более естественным, однако, является порядок, при котором название специалиста образуется на основе термина, обозначающего дисциплину. Поэтому применение этого термина является крайне нежелательным.

В последние годы как в нашей стране, так и за рубежом ощущается острая потребность в термине, который бы однозначно определял научную дисциплину, занимающуюся исследованием методов и средств сбора, переработки, хранения, поиска и распространения научной информации. Справедливо считая термин «документация» слишком узким и многозначным для обозначения этой научной дисциплины, зарубежные специалисты предлагают вместо него другие термины, основанные на использовании понятия информация. Наиболее распространенным из таких терминов следует признать термин «информационная наука» (information science). Ниже приводится одно из определений этого термина.

«Информационная наука — это отрасль науки, исследующая свойства и поведение информации; силы, управляющие потоком информации; средства обработки информации с целью обеспечения ее оптимальной доступности и использования. Изучаемые этой отраслью науки процессы включают создание, распространение, сбор, организацию, хранение, поиск, интерпретацию и использование информации. Данная отрасль науки ведет свое происхождение или связана с математикой, логикой, лингвистикой, психологией, вычислительной техникой, исследованием операций, полиграфией, связью, библиотековедением, научной орга-

---

[84] В. А. Успенский и Ю. А. Шрейдер. К проблематике теории научной информации. «Научно-техн. информация», 1963, № 3, стр. 17—20.

[85] Кибернетика и документалистика. Механизация процесса нахождения, хранения и поиска научной информации. М., изд-во «Наука», 1966.

мации», опубликованной в сборнике «Научно-техническая информация» [90].

По-видимому, пришла пора сломать неудачно складывающуюся традицию и ввести новый термин для обозначения новой научной дисциплины. Вполне приемлемое предложение такого термина содержится в упомянутой рецензии профессора Я. Г. Дорфмана. Он пишет: «Более логичным было бы назвать эту новую дисциплину, трактующую о принципах, методах и средствах сбора, переработки, хранения, поиска и распределения любого вида информации — информатикой» [91].

В связи с этим нельзя не вспомнить, что несколько лет назад подобное же предложение исходило от директора Института проблем передачи информации АН СССР, члена-корреспондента АН СССР, покойного А. А. Харкевича. В письме от 11 октября 1962 г. на имя А. И. Михайлова он писал: «Не подлежит сомнению, что проблема организации информационной деятельности в области науки становится — и уже стала — одной из крупных и острых проблем. Также бесспорно, что успешное и радикальное решение этой проблемы возможно лишь на основе научных методов и подходов, которые в большей части нужно создавать заново. Со всем, что из этого вытекает и что с необходимыми подробностью и ясностью изложено в Вашей записке [92], я полностью согласен. Можно было бы спорить о терминологии (раздел 1, 4), но, учитывая оговорку о том, что принятая терминология ограничивается задачами записки, дискуссию на эту тему можно и отложить. Единственно, что, пожалуй, не терпит отлагательства, — это вопрос о наименовании новой дисциплины. «Научная информация» — это, конечно, наименование объекта науки, а не самой науки. Я бы не остановился перед введением нового термина; пропагандировать его Вам легче, чем кому-либо. Например, «информология» или «информатика» («информация» плюс «автоматика») или т. п.».

Необходимо отметить, что и за рубежом предполагалось для обозначения новой научной дисциплины использовать термин «информатика». В 1962 г., например, французский специалист Ф. Дрейфус привел варианты этого термина на пяти языках «informatique» (франц.), «informatics» (англ.), «Informatick» (нем.), «informatica» (итал., исп.) [93].

Новый термин с четко определенным и строго фиксированным значением часто оказывается лучше старых терминов, которые,

[90] Я. Г. Дорфман. Рецензия на книгу А. И. Михайлова, А. И. Черного, Р. С. Гиляревского «Основы научной информации». М., изд-во «Наука», 1965. «Научно-техническая информация», 1966, № 7, стр. 46—47.

[91] Там же, стр. 46.

[92] Имеется в виду брошюра «Научная информация (Вопросы советской науки)». Под ред. А. И. Михайлова. (Изд. 1). М., ВИНИТИ, 1962.

[93] Ph. Dreyfus. L'informatique. «Gestion», 1962, v. 5, juin, стр. 240—241.

как правило, многозначны, употребляются в различных смыслах, что затрудняет их понимание и оказывает неблагоприятное влияние на развитие самой научной дисциплины.

Таким образом, из двух терминов, на основе которых можно образовать название новой дисциплины («документ» и «информация»), более предпочтительным является второй. Из возможных сочетаний данной основы с суффиксами и окончаниями «-ия», «-логия», «-ве́дение», «-знание», «-ика» было выбрано последнее, т. е. «информат-ика». Следует отметить, что в последние годы этот способ является наиболее продуктивным для образования обозначений новых научных дисциплин. Термины «кибернетика», «бионика», «семиотика», «космонавтика» образованы именно таким образом.

По нашему мнению, термин «информатика» является наиболее удачным для названия новой научной дисциплины, и мы предлагаем ввести его в этом качестве. В результате образуется система терминов, обозначающих основные понятия данной области знаний: «информация», «научная информация», «научно-информационная деятельность», «информатика», «информатор», «информатик». Ниже даются определения каждому из этих терминов.

*Информация* — «. . . есть сырой материал и состоит из простого собрания данных, тогда как знание предполагает некоторое размышление и рассуждение, организующее данные путем их сравнения и классификации» [94]. Такое определение понятия «информация» — разумеется, весьма ограниченное — вполне достаточно для наших целей. Здесь не делается попытки дать этому понятию философское определение, хотя в самом общем виде понятие «информация» можно было бы определить как объективное содержание связи между взаимодействующими материальными объектами, проявляющееся в изменении состояний этих объектов.

*Научная информация* — это получаемая в процессе познания логическая информация, которая адекватно отображает закономерности объективного мира и используется в общественно-исторической практике. В этом определении названы четыре наиболее существенных, по нашему мнению, признака, которые необходимы и достаточны для раскрытия понятия «научная информация».

В приведенном определении говорится, что научной может быть лишь информация, получаемая в процессе познания закономерностей объективной действительности. В связи с этим необходимо подчеркнуть, что основу процесса познания составляет практика, производственная деятельность людей. «Точка зрения жизни, практики должна быть первой и основной точкой зрения теории познания», — указывал В. И. Ленин [95]. К практике отно-

[94] Л. Бриллюэн. Наука и теория информации. М., Физматгиз, 1960, стр. 13.

[95] В. И. Ленин. Полное собрание сочинений, т. 18, стр. 145.

сится материальное производство, научный эксперимент, классовая борьба, национально-освободительное движение, т. е. все виды активной деятельности людей по преобразованию природы и общества. Практика носит ярко выраженный общественный характер: она является деятельностью не отдельных личностей, а всех трудящихся. Из сказанного следует, что источником научной информации служит не только исследовательская деятельность ученых, но и производственная, хозяйственная деятельность инженеров, рабочих-новаторов и работников сельского хозяйства, их многочисленные изобретения, усовершенствования и рационализаторские предложения.

Но не всякая информация, получаемая в процессе познания внешнего мира, есть научная информация. Чувственное познание дает человеку представление лишь о единичном, т. е. об отдельных, внешних сторонах вещей. Оно не в состоянии раскрыть общего, внутренней природы вещей, тогда как только знание сущности вещей, законов природы и общественной жизни может служить для человека руководством в его практической деятельности. «. . . задача науки, — отмечал К. Маркс, — заключается в том, чтобы видимое, лишь выступающее в явлении движение свести к действительному внутреннему движению. . .» [96]. Внутренняя природа вещей познается благодаря логическому мышлению, облеченному в языковую форму. Поэтому информация становится научной лишь тогда, когда она подвергнута обработке и обобщению абстрактно-логическим мышлением. Именно этим научная информация отличается от сведений, получаемых человеком в процессе чувственного познания.

Третий необходимый признак научной информации — адекватное отображение закономерностей объективной действительности. Под адекватным отображением в данном случае понимается обусловленная достигнутым уровнем развития науки степень правильности отображения объективной реальности. В процессе познания люди могут получать логическую информацию, которая дает искаженное представление об окружающем мире. Таковы религиозные абстракции, не имеющие никакого отношения к научной информации. В связи с этим необходимо отметить, что ошибочные абстракции — гипотезы и теории — выдвигаются и в науке. Но ошибочность таких гипотез и теорий не свидетельствует об их ненаучности. Ошибочная гипотеза или теория вполне научна, если ведется систематическое изучение, обобщение и проверка на практике положений этой гипотезы или теории. В этом состоит принципиальное отличие ошибочных научных гипотез и теорий от лженаук — астрологии, алхимии и др.

Наконец, четвертый признак научной информации состоит в том, что она непременно должна использоваться в общественно-

[96] К. Маркс и Ф. Энгельс. Сочинения, т. 25, ч. 1, стр. 343.

исторической практике. Это обеспечивает постоянную проверку истинности логической информации и препятствует отнесению к категории научной информации общеизвестных или устаревших истин, научной фантастики, предписаний народной медицины и т. д. От такого рода ненаучных знаний следует отличать чисто теоретические знания, которые в конечном итоге всегда находят практическое применение. Здесь хотелось бы подчеркнуть, что вопрос о принадлежности тех или иных знаний к категории научной информации может быть решен (и всегда должен решаться) лишь с учетом конкретно-исторических условий, в которых эти знания стали объектом нашего внимания. Разумеется, такие знания прежде всего должны обладать всеми перечисленными выше признаками, которые необходимо рассматривать в тесном диалектическом единстве друг с другом.

Еще одно замечание. В термине «научная информация» слово «научная» не означает, что такая информация является атрибутом лишь науки, научно-исследовательской деятельности. Определение «научная» употребляется в этом термине с целью подчеркнуть, что обозначаемая им информация получена в процессе познания объективного мира и адекватно отображает его закономерности. По отношению к понятию «научная информация» видовыми являются понятия «астрономическая информация», «биологическая информация», «геофизическая информация», «химическая информация» и т. д. (деление по отраслям науки, в которых получают и используют научную информацию); «техническая (или научно-техническая) информация», «политическая информация», «сельскохозяйственная информация», «управленческая информация» и т. д. (деление по областям практической деятельности людей).

*Научно-информационная деятельность* — это действия, выполняемые в целях сбора, аналитико-синтетической переработки, хранения, поиска и распространения научной информации, получаемой в процессе познания объективного мира. Научно-информационная деятельность — это самостоятельная часть научного труда, выделившаяся в ходе его общественного разделения и имеющая своей задачей обеспечение заинтересованных специалистов необходимыми им сведениями о результатах научных исследований и технических разработок, о производственном опыте и т. д.

*Информатика* — это научная дисциплина, изучающая структуру и свойства (а не конкретное содержание) научной информации, а также закономерности научно-информационной деятельности, ее теорию, историю, методику и организацию. Целью информатики является разработка оптимальных способов и средств представления (записи), сбора, аналитико-синтетической переработки, хранения, поиска и распространения научной информации. Информатика имеет дело со смысловой (семантической)

информацией, но не занимается качественной оценкой этой информации. Такая оценка может производиться лишь специалистами в соответствующих областях науки или практической деятельности.

*Информатор* — это специалист в какой-либо отрасли науки или практической деятельности, занимающийся исключительно научно-информационной деятельностью и использующий в своей работе достижения информатики.

*Информатик* — это специалист в области информатики. Кадры информатиков пополняются, главным образом, за счет наиболее квалифицированных информаторов.

Введение терминов «информатор» и «информатик» вызвано необходимостью различать по названию специалистов, занимающихся научно-информационной деятельностью, и специалистов в области информатики. В английском языке, например, для этой цели используются термины information officer (информатор) и information scientist (информатик).

Предлагаемая система терминов не является совершенно новой. Она лишь ограничивает сферу употребления термина «научная информация» ее первым и основным значением и вводит новый термин — «информатика» и производный от него термин «информатик». Тем не менее введение этой системы терминов потребует некоторого изменения существующего словоупотребления. Нам представляется, что такие изменения совершенно необходимы, так как создание и использование точной и однозначной терминологии в рассматриваемой нами отрасли науки будет способствовать дальнейшей кристаллизации новой научной дисциплины — информатики.

Любая научная дисциплина, претендующая на самостоятельность, должна отвечать ряду вполне определенных требований, важнейшие из которых — реальность предмета изучения, практическая необходимость разработки, проникновение в сущность изучаемых предметов и явлений, выявление присущих им закономерностей, многоплановость и многообъективность и т. д.

Применительно к информатике преобладающее большинство этих требований выполняется полностью. Тем не менее нельзя считать информатику уже достаточно разработанной. Она находится в начале своего становления и переходит от этапа сбора фактов и их обобщения к этапу анализа фактов и выявления основных закономерностей, присущих изучаемому предмету. Этот путь неизбежно проходит любая новая отрасль науки. Результаты, которые достигнуты новой дисциплиной за короткий период ее становления, говорят о ее необходимости и перспективности.

Основная задача теоретического плана, стоящая перед данной дисциплиной, заключается в выяснении закономерностей, в соответствии с которыми происходит создание семантической информации, ее преобразование, передача и использование в различных

сферах деятельности человека. Прикладной характер исследований в этой области заключается в разработке наиболее рациональных методов осуществления информационных процессов, в определении способов наиболее оптимальной организации связи как внутри науки, так и между наукой и производством, с широким применением современных технических средств.

Выявляя общие свойства научно-информационных процессов, информатика при разработке наиболее рациональных методов и средств передачи научной информации исходит из существующей системы отраслей науки с присущими каждой отрасли особенностями. В этом отношении можно считать, что существует практическая связь рассматриваемой дисциплины со многими другими отраслями науки и техники, которые охватываются информационной деятельностью.

Область изучения данной научной дисциплины состоит из трех основных частей.

1) информационная деятельность как совокупность процессов подготовки научных документов, их сбора, аналитико-синтетической переработки, хранения, поиска и распространения информации, содержащейся в этих документах;

2) документы как материальные носители научной информации;

3) средства обеспечения информационных процессов.

Такое разделение можно считать на данной стадии изучения вопроса оправданным. Когда выдвигается задача изучения какого-либо процесса, мы неминуемо сталкиваемся с необходимостью исследования тех объектов, которые участвуют в этом процессе, а также тех средств, которые обеспечивают осуществление данного процесса. Разумеется, изучению подлежат и причины, которые вызывают и формируют этот процесс, а также влияют на ход его развития.

В информатике научно-информационная деятельность рассматривается не только как форма связи внутри науки, но и как форма взаимной связи науки с производством и другими видами человеческой деятельности. При этом под научно-информационной деятельностью понимается не только совокупность процессов, которые выполняются в специальных информационных органах и учреждениях, в нее входят также процессы, связанные с созданием и использованием документальной информации.

Сформулированные выше в самых общих чертах предмет и задачи информатики подтверждают право на самостоятельность этой новой дисциплины. В настоящее время не представляется возможным назвать другие научные направления, перед которыми стояли бы подобные задачи и в которых исследовались бы сходные проблемы.

Если рассматривать предмет и задачи информатики в более конкретном виде, то связанные с нею научно-исследовательские

работы, ведущиеся в настоящее время, можно сгруппировать по следующим основным направлениям:

1. Изучение информационных процессов: а) сбор и классификация документальной информации; б) методы создания и разработки оптимальных форм вторичных научных документов и изданий; в) методы хранения и поиска семантической информации; г) методы распространения информации; д) организация научно-информационной деятельности в различных странах и различных областях человеческой деятельности.

2. Разработка методов анализа семантической информации (определение критериев оценки содержания документов; методы извлечения наиболее существенной информации; методы переработки информации и т. п.).

3. Разработка методов закрепления научной информации в документальном виде.

4. Разработка информационно-поисковых систем применительно к различным объемам и видам информации. Разработка машинных информационных языков и способов перевода с естественных языков на информационные и обратно.

5. Исследование возможностей применения машинной техники для обеспечения информационных процессов. Изучение применения электронных цифровых вычислительных машин как средства реализации информационно-поисковых систем.

Перечень этих направлений определяет тематику большей части работ по информатике, опубликованных в СССР и за рубежом. В этом же плане проводится аспирантская подготовка и вообще подготовка кадров информационных работников.

Следует отметить, что общее число работ по теории и практике научно-информационной деятельности составляет около 5 тыс. в год. Они публикуются более чем в 50 журналах и продолжающихся изданиях. В крупных странах насчитываются десятки тысяч специалистов, занятых в этой области; определенная их часть занимается только исследовательской работой.

Не останавливаясь далее на формах и методах научно-информационной деятельности, которые подробно рассматриваются в соответствующих главах книги, считаем необходимым в данной главе осветить лишь основное содержание научно-информационного процесса, который состоит из следующих основных этапов:

1. Сбор научных документов.
2. Аналитико-синтетическая переработка документальной информации.
3. Хранение и поиск информации.
4. Репродуцирование и распространение информационных материалов.

Мы останавливаемся на этих вопросах только в самых общих чертах, только в связи с рассмотрением предмета и проблем информации. Эти этапы научно-информационного процесса в сово-

купности представляют собой результат развития информационной практики. При этом каждый этап может рассматриваться в качестве самостоятельной достаточно обширной части этого процесса, характеризующейся специфическими методами и средствами.

До последнего времени информатика под воздействием практических потребностей более интенсивно решала отдельные частные вопросы. Общие проблемы информатики как самостоятельной отрасли знания, начиная с определения ее предмета, выявления объективных связей и закономерностей и кончая вопросами терминологии, оставались малоизученными и поэтому нуждаются в дальнейшей серьезной теоретической разработке.

## ИНФОРМАТИКА И ДРУГИЕ НАУЧНЫЕ ДИСЦИПЛИНЫ

Любая новая научная дисциплина связана с другими отраслями знания. Эти связи можно рассматривать как по общности предмета или метода исследования, так и в историческом плане. Информатика в той или иной степени связана с такими научными дисциплинами, как математическая теория информации, кибернетика, семиотика, математическая логика, логическая семантика, психология, книговедение, библиотековедение, библиография, а также с различными отраслями техники.

Успешное применение *математической теории информации*, основы которой разработаны Клодом Шенноном, в самых различных областях знания, ее тесная связь с кибернетикой породили неверное представление о всеобъемлющем характере этой важной математической дисциплины. Это обстоятельство подчеркивал и сам К. Шеннон. В 1956 г. он писал: «За последние несколько лет теория информации превратилась в своего рода бандвагон [97] от науки. Появившись на свет в качестве специального метода в теории связи, она заняла выдающееся место как в популярной, так и в научной литературе. Это можно объяснить отчасти ее связью с такими модными областями науки и техники, как кибернетика, теория автоматов, теория вычислительных машин, а отчасти новизной ее тематики. В результате всего этого значение теории информации было, возможно, преувеличено и раздуто до пределов, превышающих ее реальные достижения. Ученые различных специальностей, привлеченные поднятым шумом и перспективами новых направлений исследования, используют идеи теории информации при решении своих частных задач. Так, теория информации нашла применение в биологии, психо-

[97] Бандвагон — американский термин, означающий политическую группу или партию, программа которой находит шумную поддержку населения. — *Прим. авторов.*

логии, лингвистике, теоретической физике, экономике, теории организации производства и во многих других областях науки и техники. Короче говоря, сейчас теория информации, как модный опьяняющий напиток, кружит голову всем вокруг. . . Основу теории информации составляет одна из ветвей математики, т. е. строго дедуктивная система. Поэтому глубокое понимание математической стороны теории информации и ее практических приложений к вопросам общей теории связи является обязательным условием использования теории информации в других областях науки. Я лично полагаю, что многие положения теории информации могут оказаться очень полезными в этих науках; действительно, в ней уже достигнуты некоторые весьма значительные результаты. Однако поиск путей применения теории информации в других областях не сводится к тривиальному переносу терминов из одной области науки в другую» [98].

Иногда высказывается мнение, что информатика не более как частный случай математической теории информации или одно из ее приложений. Необходимо признать, что для подобного мнения имеются известные основания. Действительно, методы математической теории информации достаточно широко применяются и в информатике. Однако следует со всей определенностью провести четкую грань между математической теорией информации и информатикой.

Центральными понятиями математической теории информации являются понятия информации и меры для измерения ее количества. В этой теории определение информации выводится из статистических рассуждений и полностью игнорирует смысловую сторону передаваемых сообщений. На это неоднократно указывали ведущие специалисты в области математической теории информации. Л. Бриллюэн в своей книге «Наука и теория информации» писал: «Мы определяем информацию как результат выбора; мы не рассматриваем информацию как основание для предсказания, как результат, который может быть использован для того, чтобы сделать другой выбор, мы полностью игнорируем человеческую оценку информации. . . В соответствии с нашим определением совокупность 100 букв, выбранных случайным образом. . ., фраза в 100 букв из газеты, пьесы Шекспира или теоремы Эйнштейна имеют в точности одинаковое количество информации. Другими словами, мы определяем информацию как нечто отличное от знания, для которого у нас нет количественной меры» [99]. И далее: «Только игнорируя человеческую ценность информации, мы смогли построить научную теорию информации, основанную на статистике, и была показана полезность этой

[98] К. Шеннон. Бандвагон. В кн.: К. Шеннон. Работы по теории информации и кибернетике. М., ИЛ, 1963, стр. 667—668.

[99] Л. Бриллюэн. Наука и теория информации. М., Физматгиз, 1960, стр. 29—30.

теории. Существует, однако, много проблем, которые нельзя разрешить методами этой теории» [100].

Об этом же писал и С. Голдман: «. . .количественная теория информации может показаться неполной и даже разочаровать читателя, ибо в ней ничего не говорится о ценности информации. Сообщение о том, родила ли жена Джона Смита мальчика или девочку, содержит столько же двоичных единиц информации, сколько и сообщение о том, кого родила ваша жена. Вместе с тем последнее сообщение представляет для вас неизмеримо бóльшую ценность, чем первое. Может быть, и удастся построить математическую теорию, которая установит соотношения между информацией и ее потребителем и таким образом позволит оценивать в цифрах не только количество, но и ценность информации. Насколько известно автору, в настоящее время такая теория отсутствует» [101].

Итак, математическая теория информации целиком и полностью отвлекается от содержательной, семантической стороны информации, тогда как в информатике именно эта сторона является решающей. Поэтому математическая теория информации не может охватить всех аспектов информатики и не может рассматриваться как ее математическая основа.

Методы математической теории информации могут быть использованы и уже используются в информатике для обеспечения оптимального кодирования семантической информации, вводимой в электронные цифровые вычислительные машины, для ее долговременного хранения, поиска, передачи на расстояние и выдачи, при переработке текстов документов на естественных языках, например для сокращения их объема без потери смысловой информации и т. д.

В математической теории информации одной из основных задач является рассмотрение процесса передачи информации по каналу связи. Если в информатике рассматривать информационную систему (в широком смысле) как своеобразный канал связи, то можно увидеть определенное сходство в механизме передачи информации в том и другом случае. Один из признаков такого сходства между системами для передачи сообщений и системами для хранения и поиска информации состоит в том, что обе системы обладают четко выраженным свойством адресности.

При передаче сообщений всегда соблюдается, по крайней мере, одно из двух условий:

а) точно известны пространственные координаты (адрес) индивидуального или коллективного потребителя информации;

б) индивидуальному или коллективному потребителю информации точно известны параметры канала связи, по которому

---

[100] Там же, стр. 379.

[101] С. Голдман. Теория информации. М., ИЛ, 1957, стр. 76—77.

передается нужная ему информация (например, несущая частота и время передачи).

Системы для хранения и поиска информации также обладают свойством адресности: авторы научных документов — книг, статей, научно-технических отчетов и т. п. — знают, кому адресуется сообщаемая ими информация. Материальное закрепление научной информации в форме печатного документа означает, что эта информация рассчитана на коллективного потребителя и на сохранение во времени.

Представляется вполне допустимым, что одной из причин появления периодических изданий была потребность внести большую степень адресности в систему передачи научной информации при помощи публикаций, а также сократить время, затрачиваемое на передачу этой информации. Поскольку периодические издания имеют профилированный характер, помещаемая в них научная информация приобретает свойство определенной адресности. Если статья помещена, например, в журнале по нефтехимии, то это означает, что она адресована уже не вообще всем химикам, а лишь тем из них, которые работают в этой области химии. При этом адресность публикаций повышается по мере углубления специализации периодических изданий. Можно предположить, что именно необходимость повысить адресность научной информации является одной из причин того необычайно быстрого роста числа периодических изданий, который наблюдается в настоящее время.

Завершая рассмотрение вопроса о связях информатики с математической теорией информации, необходимо отметить следующее обстоятельство. В последнее время предпринимаются попытки построить на базе теории информации К. Шеннона теорию семантической информации, учитывающую не только количество, но и ценность передаваемых сообщений. Наиболее известной в этом направлении является работа Р. Карнапа и И. Бар-Хиллела [102]. Если эти попытки увенчаются успехом, то такая теория найдет самое широкое применение в информатике.

Более сложны связи информатики с *кибернетикой*. Некоторые ученые считают, что информатика есть часть кибернетики. Академик А. Н. Колмогоров определяет кибернетику «как науку о способах восприятия, хранения, переработки и использовании информации в машинах, живых организмах и их объединениях» [103]. Академик А. И. Берг включает в прикладные лингвистические проблемы современной кибернетики автоматизацию информационной службы, перевода и реферирования научно-технической

[102] R. Carnap, J. Bar-Hillel. An outline of a theory of semantic information. «British Journal of Philosophical Science», 1953, v. 4, стр. 147—157.

[103] А. Н. Колмогоров. Кибернетика. «БСЭ», изд. 2, т. 51, М., 1958, стр. 149.

литературы, построение информационно-логических машин, конструирование стенографов-автоматов и другие важные задачи информатики [104]. По этому поводу трудно высказать определенное мнение, так как сама кибернетика формируется на стыке многих наук и границы ее весьма неопределенны. Не закончилось и формирование информатики, которую авторы считают самостоятельной научной дисциплиной. Окончательный ответ на этот вопрос может быть получен лишь на более поздней стадии развития этих научных дисциплин.

В информатике широко используются также методы *семиотики*, математической логики и логической семантики. Любая информация всегда записывается на том или ином конкретном языке. Это может быть естественный язык с определенной научной терминологией и специальными символами или же искусственный язык типа языков программирования. При машинной записи и обработке семантической информации используются формализованные информационные языки. Поэтому для информатики чрезвычайно важно изучение различных языков, пригодных для записи информации, а также построение таких языков. В этой части информатика примыкает к семиотике — общей теории знаковых систем.

Необходимо также изучать средства записи информации в естественных языках и способы перехода от одних форм записи к другим. Существенным является разработка методов, позволяющих анализировать смысл текстов, записанных на различных языках.

Наряду с исследованиями языков необходимо разрабатывать алгоритмы для преобразования информации как в пределах одного языка, так и при переходе с одного языка на другой.

На первом этапе создания информационных языков и алгоритмов необходимо проведение соответствующих метатеоретических исследований. Создавая информационный язык для любой научной дисциплины, необходимо прежде всего уяснить содержание основных понятий, используемых в этой дисциплине, способы их записи, их взаимоотношения и связи. В этом случае к области метатеоретических исследований относится установление способов умозаключений, применяемых в различных отраслях науки, а также определение «существенности» того или иного факта. Это необходимо для разработки методов получения новой информации путем логической переработки уже имеющейся информации, а также для разработки алгоритмов реферирования.

В комплекс метатеоретических исследований входит изучение средств записи информации, применяемых в данной науке,

---

[104] А. И. Берг. Предисловие к русскому изданию. В кн.: М. Таубе. Вычислительные машины и здравый смысл. М., 1964, стр. 9.

и разработка ее формального аппарата. Именно разработанность формального аппарата математики позволила сразу применить новую технику для решения вычислительных задач математики. Соответствующая формализация языков других наук позволит аналогичным образом применить новую технику для решения их специфических задач. Изучение существующих и создание новых способов записи информатики требует решения ряда сложных теоретических вопросов современной *логической семантики*.

Методы *математической логики* используются в информатике как для непосредственного создания информационных языков, так и для формализации процессов логического вывода в тех или иных теориях. Эта формализация в настоящее время разработана лишь для математики. Распространение подобной формализации на другие теории является важной общенаучной задачей. Математическая логика обслуживает главным образом математику.

В связи с этим необходимо создание аналогичных прикладных «логик», обслуживающих другие науки. Работы в этом направлении уже ведутся в области химии. К кругу логико-математических исследований относятся также исследования, касающиеся дедуктивного построения той или иной области науки, выяснения принципиальных возможностей формализованного и, следовательно, машинного оперирования с предложениями в этой области. Необходимо разрабатывать также методику программирования информационных алгоритмов.

Особый интерес для информатики представляет изучение возможных методов формализованного реферирования, а также проблемы машинного перевода, основы которого разрабатываются психологией мышления, лингвистикой и семиотикой.

Эти проблемы могут быть наилучшим образом решены лишь при условии проведения глубоких исследований в области теории и методики аннотирования, реферирования и научного перевода, а также большой экспериментальной работы. Важное значение для реализации машинного перевода имеет решение проблем, связанных с семантическими аспектами языка и разработкой терминологии.

Исключительный интерес для информатики представляют некоторые методы современной *психологии*. До последнего времени основные усилия службы научной информации были направлены на всемерное сокращение сроков подготовки различных информационных изданий — реферативных журналов, библиографических указателей, каталогов и т. п. При этом почти не делалось никаких попыток вмешаться в процесс создания первичных научных документов: монографий, статей, научно-технических отчетов и т. п. Однако задачи службы научной информации могут решаться значительно эффективнее, если повысить контроль за качеством рукописей, отобранных для опубликования.

Было бы глубоким заблуждением считать эти задачи чисто организационными. Научно обоснованный отбор рукописей для опубликования можно производить лишь тогда, когда разработаны объективные критерии определения научной ценности этих рукописей. В процессе редактирования рукописям должна быть придана оптимальная форма. Для решения этих задач необходимо глубокое исследование психологии творческого труда ученого и инженера, изучение его потребностей в научной информации на разных этапах научно-исследовательской или опытно-конструкторской работы, а также специфики документа как важнейшего носителя научной информации. Нужно точно установить, какая научная информация требуется исследователю на каждой стадии его работы, в каком количестве и в какой форме. Необходимо разработать типологию научных документов и оптимальную форму каждого типа. Ответы на эти вопросы можно получить лишь при широком использовании различных методов современной психологии.

Чтение научно-технической литературы есть органическая часть любого научного исследования. Поэтому научно-информационная деятельность не преследует цели предельно разгрузить или даже совсем освободить исследователя от такого чтения. Служба научной информации должна не только обеспечить научного работника или инженера документальной информацией, имеющей оптимальную форму и поступающей к нему в нужный момент и в нужном количестве, но и научить его эффективно использовать эту информацию, привить ему навыки быстрого чтения. Эти, а также многие другие научно-информационные задачи (например, изучение систем класса «человек—машина», обеспечение высокой четкости печатных шрифтов, отображение знаков и т. п.) решаются методами психологии.

Алгоритмизация процессов реферирования, аннотирования и индексирования возможна лишь при условии понимания нами сущности процессов решения этих задач человеком. Для этого необходимо глубокое изучение механизмов человеческого мышления. Ключевыми проблемами информационного поиска (помимо индексирования) являются алгоритмическое адекватное выражение информационной потребности на том или ином естественном языке, а также алгоритмическое определение смысловой эквивалентности двух текстов.

Эти проблемы могут быть решены лишь в результате глубоких исследований в области психологии мышления.

Развитие человеческого общества, его науки, техники и культуры невозможно без накопления семантической информации и ее передачи от одного поколения людей к другому. В этом процессе важную роль играют способы передачи информации. История человечества свидетельствует о том, что способы передачи знаний оказывают огромное влияние на темпы общественного прогресса.

«Если рассматривать историю науки и техники в связи с уровнем средств передачи и хранения информации, то сравнительно четко выделяются пока три большие эпохи: первая и самая продолжительная (когда основным средством передачи информации была почти исключительно устная речь); вторая — от зарождения классового общества и основания рабовладельческих государств — характеризуется. . . добавлением к устной речи письменности; и, наконец, третья эпоха — примерно от становления буржуазного общества по сей день — в которой основным средством хранения и передачи научной и технической информации становится книгопечатание» [105].

Действительно, эпоха первобытнообщинного строя, для которой характерен наиболее примитивный способ передачи семантической информации — устная речь, длилась несколько сотен тысячелетий. В эпохи рабовладельческого строя и феодализма основным средством хранения и передачи человеческих знаний была письменность — значительно более совершенный способ по сравнению с устной речью. Эта эпоха длилась несколько тысяч лет. В середине XV в. появился наиболее совершенный из всех до сих пор применявшихся способов передачи информации — книгопечатание. С тех пор прошло всего пятьсот лет, но за этот исторически ничтожный промежуток времени сменились две общественно-экономические формации, появилась и успешно развивается мировая социалистическая система, и развитие науки, техники и культуры достигло больших высот. Поскольку способы хранения и передачи семантической информации играют такую большую роль в развитии человеческого общества, небезынтересно проследить эволюцию этих способов.

На определенном этапе развития человеческого общества устная речь уже не могла справиться с коммуникативной функцией. Потребность передавать информацию на расстояния, значительно превышающие возможности человеческого голоса, а также необходимость передавать информацию во времени поставили человечество перед необходимостью изыскивать средства материализации речи, закрепления сообщаемой информации на продолжительное время. Возникает письменность — система материального закрепления звуков речи и слов языка при помощи соответствующих условных знаков с целью передачи их в пространстве и времени.

*Книговедение* и, в частности, история книги дает ценные сведения относительно важнейших этапов формирования научных документов. С развитием каждой цивилизации, по мере усовершенствования языка и его письменности, вырабатывались и опре-

---

[105] О. С. Кулагина, Л. А. Калужнин, В. В. Иванов. О культурном значении машинного перевода. «Вестник истории мировой культуры», 1961, № 3 (27), стр. 23.

деленные типы документов, произведений с законченной формой и минимумом библиографических сведений. Эти документы, наряду с устным общением, служили для передачи знаний об окружающем мире и навыков, используемых в том или ином ремесле. Самые древние из дошедших до нас научных документов — клинописные плитки Месопотамии — датируются четвертым тысячелетием до н. э. От шумерской эпохи (четвертое тысячелетие до н. э.) и вавилоно-ассирийской культуры (расцвет во втором тысячелетии до н. э.) сохранились сочинения в области истории, права, астрономии, географии, торговли. Наряду с небольшими глиняными плитками высотой в 2,5 *см* изготовлялись плитки высотой до 40 *см*, содержащие до 400 строк в 4 колонки с двух сторон.

Основным недостатком этих «глиняных книг», в течение тысячелетий обслуживавших культурные и научные потребности человека, была их громоздкость и недостаточная емкость. Отдельные произведения занимали до 10 плиток. Собрания, насчитывающие несколько десятков тысяч плиток (до нас дошли остатки храмовых библиотек в Ниппуре, датируемых третьим и вторым тысячелетием до н. э., и библиотека ассирийского царя Ашшурбанипала в Ниневии, относящаяся к 7 в. до н. э.), требовали огромных помещений.

Папирусный свиток — более компактная форма документа, позволявшая накапливать большие собрания произведений письменности. Начиная с третьего тысячелетия до н. э. в Египте изготовлялись именно такие книги. Текст в папирусном свитке располагался перпендикулярно его длине колонками от 25 до 45 строк. Хрупкость и недолговечность папируса обусловили незначительное число дошедших до нас образцов древнеегипетских документов (древнейший из них восходит к XVIII в. до н. э.). Это — главным образом ритуальные «книги мертвых», извлекаемые археологами из пирамид и других захоронений. У греков и римлян на протяжении долгого времени научные сведения распространялись устным путем. Известно, что лишь после Аристотеля для этой цели стали применяться рукописи. Ни одного оригинала времен античных классиков не сохранилось, часто столетия отделяют последнюю сохранившуюся копию от возможной даты написания текста. Сведения о греческих и римских папирусах почерпнуты из малоазиатских и позднеегипетских образцов. Средняя длина свитка не превышала 10 *м*, ширина 30 *см*.

Пергамент, изготовлявшийся из телячьих шкур, известен как материал для письма с III в. до н. э. Он позволял писать с двух сторон, был более долговечным, чем папирус, и обусловил переход к современной блочной форме книги — кодексу. Кодексы из папируса изготовлялись еще в первые века до н. э., но к V в. были вытеснены пергаментным кодексом. Кодекс — более емкая форма книги, чем свиток, он удобнее для записи больших текстов и

для наведения справок. Пергаментные кодексы вплоть до XII в. были единственной формой книги в Европе. Основным их недостатком, препятствовавшим широкому распространению письменных документов, была дороговизна. Для одного экземпляра пергаментной книги требовались шкуры целого стада.

На смену пергаменту пришла бумага, изобретенная в Китае в 105 г.; на Ближнем Востоке ее начали изготовлять с 751 г., а в XII в. через арабских завоевателей Испании она проникла в Европу. Бумажная книга, значительно более дешевая, чем пергаментная, стала широко использоваться в научных и образовательных целях.

Опыт нескольких тысячелетий развития письменных документов показывает, что их форма менялась главным образом под влиянием потребностей общества: документы становились все более емкими, удобными для использования, и дешевым средством закрепления информации. Это привело к появлению бумажного книжного блока (кодекса), который до нашего времени остается основной формой документов.

В XV в. стала повсеместно ощущаться потребность в новом способе изготовления документов. Рукописная книга не удовлетворяла культурные и научные потребности общества по двум основным причинам: 1) книга изготовлялась слишком долго и требовала больших затрат труда; 2) переписка текстов от руки не давала возможности получить большое число экземпляров идентичного содержания: копии одного и того же текста отличались друг от друга из-за искажений, вносимых переписчиками. Именно к этому времени (1448 г.) относится историческое изобретение И. Гутенбергом книгопечатания подвижными литерами. Технические средства, которые легли в основу книгопечатания, — граверная и литейная техника, винодельческий пресс, преобразованный Гутенбергом в печатный станок, — были известны еще в античные времена. Однако лишь настоятельная потребность общества в быстром и точном механическом воспроизведении текстов вызвала к жизни это изобретение, совершенству которого мы и сейчас не перестаем удивляться и основными принципами которого мы пользуется для изготовления печатных документов.

«Когда после темной ночи средневековья, — писал Ф. Энгельс в «Диалектике природы», — вдруг вновь возрождаются с неожиданной силой науки, начинающие развиваться с чудесной быстротой, то этим чудом мы опять-таки обязаны производству» [106]. Далее в числе факторов, которые привели к такому быстрому развитию науки, Ф. Энгельс называет появление печатного станка.

К середине XVII в. число издаваемых научных книг настолько возросло, что следить за ними стало трудно и возникла потреб-

[106] Ф. Энгельс. Диалектика природы. М., Госполитиздат, 1955, стр. 145.

ность в новых средствах распространения научных знаний. Рост числа книг, усиление интенсивности научных исследований, расширение круга людей, интересующихся прогрессом в науке, вызвали к жизни новую форму публикации — научный журнал, который первоначально был органом текущей информации о новых книгах.

Первым научным журналом, как уже было сказано, явился «Journal des sçavants» («Журнал ученых»), который начал выходить 5.I 1665 г., т. е. еще до официального создания Французской Академии наук, органом которой он служит вплоть до наших дней. Полтора столетия страницы научных журналов были почти целиком заполнены сообщениями о новых научных книгах с кратким изложением их содержания или оценкой их достоинств. Любопытно, что для раскрытия содержания книг в то время широко пользовались прямыми заимствованиями и цитатами из текста. При этом научная хроника вначале играла второстепенную роль. Но затем в журналах по естественным наукам стали все большее место занимать сообщения о вновь открытых явлениях природы и проводимых экспериментах. Таким образом, можно считать, что уже на заре своего существования научный журнал носил реферативный характер.

Появление современных реферативных журналов связано с бурным развитием естественных наук в XIX в. и увеличением числа научных журналов.

Особенно большие успехи были достигнуты в области химии, развитие которой сопровождалось быстрым ростом специальной литературы. Именно поэтому первый реферативный журнал «Pharmaceutisches Centralblatt» появился в области химии.

За прошедшие три столетия наука проделала гигантский путь и достигла такого уровня, когда темпы ее развития и количество публикуемой научно-технической литературы сделали научный журнал — основное средство распространения научных знаний на протяжении этого периода — недостаточно эффективным. В качестве основных недостатков научных журналов в этом отношении называют: 1) их многочисленность и неразграниченность по содержанию, затрудняющие ознакомление с публикуемыми в них материалами; 2) недостаточную быстроту опубликования и, следовательно, запаздывание информации; 3) лимитированный объем статей, не позволяющий сообщать результаты исследований со всеми необходимыми подробностями; 4) дороговизну изданий [107].

В течение последних 30 лет выдвигаются различные проекты замены научных журналов другими средствами распространения знаний. В основе этих проектов лежит предложение использовать в качестве такого средства отдельные статьи, не публикуемые

[107] Р. Х. Фелпс, П. Херлинг. Средства замены научных журналов. «Бюлл. ЮНЕСКО для библиотек», 1960, т. 14, № 2, стр. 65—81.

в периодических изданиях, депонируемые в специальных отраслевых центрах с обязательной о них информацией в реферативных изданиях. Кстати говоря, один из первых проектов такого рода был выдвинут советскими учеными, делегатами Международного геологического конгресса в 1933 г.[108] В этом же году было видвинуто предложение Уотсона Дэвиса, которое легло в основу широко известного плана Джона Бернала [109].

Таким образом, сложившаяся к настоящему времени система научных документов формировалась под влиянием общественных потребностей, в первую очередь, потребностей науки. Эта система служит одним из основных предметов исследования теории научной информации и одним из главных объектов информационной деятельности. Все основные научные сведения, идеи и факты закрепляются для передачи во времени и пространстве в форме документов. Именно поэтому группа ведущих ученых США — членов Консультативного комитета по научным вопросам при президенте — в докладе «Наука, правительство и информация» писала в 1963 г.: «Идеи и фактические данные, которые составляют сущность науки и техники, отражаются в литературе; только при сохранении единства литературы возможны единство и жизнеспособность науки» [110].

Очевидно, что способ передачи научной информации определяется уровнем развития науки и техники. С другой стороны, в зависимости от того, насколько этот способ соответствует потребностям науки и техники, он может препятствовать или, наоборот, содействовать их развитию.

По общности предмета исследования и некоторых видов практической деятельности информатика связана также с *библиотечным делом и библиографией*. До возникновения информатики как самостоятельной отрасли некоторые ее задачи решались именно в этих областях. Более того, многие библиотеки и библиографические учреждения продолжают вести информационную работу в настоящее время. Это породило ряд неверных представлений о направлении дальнейшего развития библиотечного дела, библиографии и о взаимосвязях их с информатикой.

На страницах специальных периодических изданий неоднократно высказывались мнения, что на смену научно-техническим библиотекам и библиографическим учреждениям должны прийти информационные центры. Отсюда как будто бы вытекает, что

---

[108] M. I. Goldman. Reform in the system of scientific publication. «Science», 1934, v. 80 (Oct. 26), стр. 380—381.

[109] W. Davis. Project for scientific publication and bibliography. В кн.: J. D. Bernal. The social function of science. London, 1939, стр. 449—455.

[110] Science, government, and information. A report of the President's Science Advisory Committee. Washington, U. S. Government Printing Office, 1963, стр. 7.

такие дисциплины, как библиотековедение и библиография, поглотятся информатикой.

В последние годы на страницах буржуазной библиотечной прессы появились статьи, в которых утверждается, что современная библиотека переживает глубокий кризис и должна отмереть. В них высказывается достаточно обоснованная мысль, что традиционная форма книги слишком громоздка, неудобна для хранения и распространения; изготовление ее занимает много времени и требует больших материальных затрат. В качестве средств, устраняющих эти недостатки, называют микроиздания — микрофильмы и микрокарты. При этом не учитывают, что микроиздания удобны только для хранения документов, а для использования их необходимо воспроизводить или репродуцировать в натуральную величину. Отсюда делается ошибочный вывод, что в дальнейшем библиотека будет работать без книг.

Подвергается сомнению и сама библиотечная профессия. По мнению некоторых буржуазных библиотековедов, основными сотрудниками библиотек должны стать инженеры и техники, хорошо знакомые с эксплуатацией электронных цифровых вычислительных машин, аппаратов для репродуцирования и другого машинного оборудования, а также математики, программисты. В квалификации библиотекарей хорошее знание произведений печати, владение методами комплектования, организации и отражения библиотечных фондов для читателей, навыки воспитательной работы и руководства чтением считаются этими буржуазными библиотековедами мало существенными. Эти взгляды приводят их к глубоко ошибочному выводу, что библиотека в дальнейшем может работать без библиотекарей.

Большинство буржуазных авторов формулируют основные задачи библиотеки как чисто технические и утверждают, что деятельность библиотеки ограничивается приобретением произведений печати и организацией библиотечных фондов с целью их использования читателем. Отказываясь на словах от активной работы с читателем, от пропаганды книги и руководства чтением, а на деле осуществляя эту деятельность в реакционных целях правящих классов капиталистического общества, многие теоретики буржуазного библиотековедения не видят путей дальнейшего развития библиотеки. Сводя деятельность библиотеки к чисто техническим задачам, они приходят к абсурдным выводам о библиотеке без книг и без библиотекарей, не понимают разницы между задачами научно-информационной службы и библиотечного дела. В этом сказывается ограниченность буржуазной идеологии, ведущей к узкому, объективистскому взгляду на задачи библиотечной работы.

Мы исходим из того, что советская библиотека — это культурно-просветительное и научное учреждение, организующее общественное пользование произведениями печати и письменности

и призванное помогать читателю в выборе книг, руководить чтением для того, чтобы широко пропагандировать идеи коммунизма, поднимать идейно-политический и культурный уровень советского народа, мобилизовывать его на решение политических, научных и хозяйственных задач.

Центральной фигурой советской библиотеки является библиотекарь, идейно целеустремленный, высокообразованный, владеющий знанием психологии читателя, обладающий педагогическим и профессиональным мастерством, хорошо ориентирующийся во многообразии книжной продукции, умеющий пользоваться всеми новейшими техническими средствами.

Информатика разрабатывает методы решения двух основных практических задач информационной службы: быстрого оповещения о новых фактах и идеях и долговременного хранения информации с целью оперативной выдачи сведений в ответ на поступающий запрос. Поскольку в настоящее время основные факты и идеи, находящиеся в научном обращении, содержатся в произведениях печати — книгах, журналах, продолжающихся изданиях, патентных описаниях, стандартах, технических каталогах и т. п., информационная служба включает в процесс своей деятельности библиографические функции: ответы на запросы о том, что содержится в тех или иных публикациях, в каких публикациях содержится информация по данной теме или предмету. До сих пор эти задачи решались библиотекой, но они никогда не были для библиотеки единственными или даже главными. Решая эти задачи в плане общественного пользования произведениями печати, библиотека рассматривает их как частные.

С другой стороны, потребности науки и техники не обслуживаются в должной мере произведениями печати. Публикации сильно отстают от их быстрого развития. Традиционные формы публикаций — монография, статья в журнале или сборнике и т. д. — часто ограничивают возможность быстрой и точной передачи результатов эксперимента или исследования. Именно поэтому многие специалисты предлагают отказаться от привычной системы публикаций. Уже сейчас в ряде областей науки такие неопубликованные документы, как диссертации, лабораторные журналы, отчеты об исследовательских и опытно-конструкторских работах, отчеты о научных командировках, протоколы научных конференций и совещаний и другие, — имеют не меньшее значение, чем публикации. Вполне возможно, что с течением времени книга в широком смысле уступит значительную часть своих научных функций другим, более совершенным средствам хранения и передачи информации. Информатика ставит перед собой также задачу логической переработки идей и фактов с целью получения новой информации.

По всей вероятности, среди этих средств большее значение получат те, которые позволят непосредственно, минуя литератур-

ную форму произведения (опубликованного или неопубликованного), отвечать на фактографические запросы. Ведь, в конечном счете, для научной информации важнее всего отвечать именно на фактографические запросы: какими свойствами обладает данный объект, какие объекты обладают данными свойствами, что известно по данной теме или о данном объекте и т. д.

Ответы на подобного рода вопросы никогда не входили в задачи библиотеки. И если справочно-библиографическим службам библиотек приходится сталкиваться с ними, то лишь потому, что в настоящее время ответы на них можно дать, используя документальную информацию, в первую очередь произведения печати.

Эти общие положения определяют существенные различия в задачах, стоящих перед органами научно-технической информации и научно-техническими библиотеками. Работники информационного центра, в отличие от библиотекарей, должны быть прежде всего специалистами в той или иной отрасли знания, связанными с разработкой определенной научной или технической проблемы. Они призваны обеспечить ученых и инженеров идеями и фактами, независимо от того, содержатся эти идеи и факты в опубликованных или неопубликованных документах или должны быть получены из этих документов логическим путем.

Вот по какому направлению может быть произведено размежевание функций библиотек и органов информации — двух институтов современного общества, созданных для удовлетворения различных его потребностей. Нам представляется несущественным, что в настоящее время в научно-информационной деятельности используются некоторые библиотечно-библиографические методы. Несомненно, что в недалеком будущем библиотечное дело и библиография также позаимствуют некоторые методы и эффективные технические средства информатики. Главное состоит в том, что основные задачи у этих дисциплин разные.

Особого внимания заслуживает вопрос о связях между информатикой и *библиографией*. «Библиография — область знания и практической деятельности, имеющая своим объектом: учет, описание, систематизацию и качественный анализ произведений печати и составление различных пособий, которые облегчают ориентировку в имеющейся литературе, пропагандируют ее и содействуют лучшему ее использованию» [111].

Библиография в течение многих веков обслуживала науку и технику. Но сейчас для нее становится все труднее решать эти задачи. Многие виды произведений печати (труды и материалы конференций и совещаний, специальные виды технической литературы, ведомственные издания, переводы и т. д.) далеко не полностью отражаются в библиографических изданиях. Еще более

[111] Е. И. Шамурин. Словарь книговедческих терминов. М., 1961, стр. 29.

серьезным является то обстоятельство, что соответствующие библиографические издания из-за несовершенства традиционных систем классификации не позволяют вести многоаспектный поиск литературы. Поэтому при пользовании библиографическим указателем неизбежно теряется значительная часть информации даже о тех произведениях, которые в нем учтены. Вместе с тем обслуживание науки и техники сведениями о произведениях печати [112] продолжает оставаться одной из важных задач библиографии.

Однако ученому или инженеру нужны не эти сведения (хотя за неимением других он до последнего времени удовлетворялся и ими) и даже не сами произведения, а идеи и факты, которые в них содержатся, и даже еще больше — идеи и факты, которых нет в этих произведениях, но которые могут быть извлечены из большого числа таких произведений логическим путем. Это уже не библиографическая задача, а задача соответствующей отрасли науки. При решении этой задачи используются определенные методы информатики.

Интересы библиографии выходят далеко за пределы науки и фактически объемлют все стороны человеческой деятельности, связанные с произведениями печати и письменности. Библиография в конечном счете стремится к пропаганде, распространению и наилучшему использованию этих произведений, преследуя при этом воспитательные, образовательные и научно-практические цели. Научно-информационная деятельность концентрируется в сфере науки и тех областей, где используются ее достижения (техники, сельского хозяйства, медицины, экономики и т. п.). Она оперирует не столько сведениями о произведениях, сколько идеями и фактами, независимо от того, были ли они опубликованы, и вообще независимо от способа и формы их фиксации во времени.

С *техническими науками* информатика соприкасается при создании средств реализации информационных систем. Важной задачей информатики является разработка методов хранения и поиска документов. В настоящее время наиболее перспективным методом хранения документальной информации является ее микрофотокодирование. Поэтому большое значение имеет разработка целого комплекса сложных технических задач, включающих создание высокоразрешающих оптических систем и специальных светочувствительных материалов.

---

[112] В данном случае *произведением* мы считаем результаты целенаправленной человеческой деятельности (факты, идеи, образы), имеющие определенную логическую взаимосвязь и завершенность и изложенные в письменном документе. *Произведение печати*, кроме того, получает обязательную редакционную апробацию и имеет четкий библиографический адрес, состоящий из стандартного набора выходных сведений (автор, заглавие, выходные данные или источник публикации).

Очень важно также создать долговременные запоминающие устройства (ДЗУ) двух типов:

а) ДЗУ с практически неограниченной емкостью, сравнительно медленным поиском и автоматической системой записи и считывания информации; б) ДЗУ небольшой емкости (порядка $10^7$ двоичных единиц) со сравнительно быстрым поиском, которое должно допускать также возможность перезаписи информации на тех же носителях.

Для создания ДЗУ большой емкости необходимы оптимальные носители информации, позволяющие при сравнительно малых объемах записывать возможно большее количество информации (высокоразрешающие фотослои, магнитные и термопластические пленки, фотохроматические материалы и др.). Выбранные носители должны обладать способностью сохранять информацию в течение неограниченно долгого времени, без дополнительной затраты энергии на перезапись, а также потреблять мало энергии на автоматическую запись и считывание.

Необходимо разрабатывать наиболее эффективные и экономичные автоматические системы записи информации в ДЗУ. Особенно большое значение имеет создание читающих устройств, обеспечивающих ввод информации непосредственно с печатного текста, а также устройства для ввода информации в ДЗУ путем непосредственного восприятия звуковой речи.

Наряду с адресными запоминающими устройствами (ЗУ) необходимо разрабатывать также устройства ассоциативного (безадресного) типа. В этом случае поиск информации должен производиться путем непосредственного обращения к самим элементам запоминающего устройства на основе информации, содержащейся в запросе.

Важное значение имеет разработка оперативных ЗУ сравнительно небольшой емкости (до $10^7$ двоичных единиц). Такие ЗУ необходимы также для решения задач переработки информации в информационно-логических системах. Скорость поиска при этом должна быть равна тактовой скорости работы соответствующей машины.

В настоящее время результаты обработки документальной информации, как правило, в конечном итоге воплощаются в ту или иную форму полиграфического воспроизведения. Поэтому оперативность печати в значительной степени определяет эффективность работы всей информационной системы.

Методы «классической полиграфии», применяемые в настоящее время для этих целей, с их сложным, длительным и трудоемким технологическим циклом, насыщенным многочисленными ручными операциями, в большинстве случаев не решают поставленных задач.

В области тиражного репродуцирования необходимо разрабатывать эффективные методы и средства:

1) быстрого набора текста;
2) механизированного изготовления оригинала-макета;
3) быстрого изготовления печатных форм.

В области разового копирования основной задачей является все более широкое использование новых методов формирования изображений (электрографических, диазографических, термографических и др.).

Быстрое развитие информационной техники позволяет ставить вопрос о создании «читающе-печатающих» автоматов, осуществляющих репродуцирование документов на основе непосредственного автоматического чтения или просмотра по заданной программе печатных или рукописных материалов. Особый интерес представляет проблема соединения читающих и печатающих устройств для автоматического набора текста непосредственно с рукописи.

Современный уровень техники делает реальной постановку задачи создания печатающих устройств, осуществляющих набор текста путем непосредственного восприятия звуковой речи.

Особо важной технической задачей является создание быстродействующих печатающих устройств к электронным информационным системам. Для ускорения прохождения потоков документов внутри информационной службы, более четкой и оперативной работы информационных центров чрезвычайно важно разрабатывать системы быстрой дистанционной передачи информации. Это требует широкого внедрения в информационную службу телевидения, современной электронной телеграфии и других средств связи в сочетании с современными средствами воспроизведения документов.

Во взаимосвязях информатики с техническими науками имеется одно обстоятельство, которое заслуживает особого рассмотрения. На первых этапах формирования информатики как самостоятельной научной дисциплины существовало довольно широко распространенное мнение, что основные трудности научно-информационной деятельности, особенно в области хранения и поиска информации, объясняются отсутствием соответствующих технических средств. Поэтому главные усилия направлялись не на разработку теоретических основ информатики, не на изучение внутренних закономерностей научно-информационных процессов, а на конструирование специальных машин и механизмов. Особенно четко такая тенденция проявилась в США, где был создан ряд весьма совершенных информационных устройств (например, информационно-поисковые системы «Миникард», «Файлсерч», «Магнавю» и др.). Однако этот путь не привел к сколько-нибудь существенному решению наиболее острых проблем информатики. К концу 50-х годов ошибочность такого подхода к решению этих трудных задач стала ясной и американским специалистам, которые начали переносить центр тяжести в своей

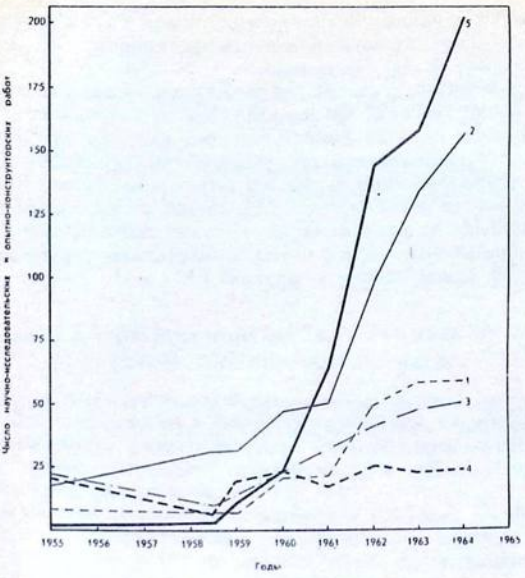



Рис. 12. Рост числа научно-исследовательских и опытно-конструкторских работ в области информатики, учтенных в издании «Current research and development in scientific documentation»

*1* — изучение информационных потребностей; *2* — исследования в области хранения и поиска информации; *3* — работы в области машинного перевода; *4* — разработка технических средств информатики; *5* — исследования в области лингвистики, анализа и синтеза речи, психологии, моделирования мыслительных процессов, машинного распознавания образов и т. п.

работе из сферы опытно-конструкторских разработок в сферу теоретических исследований. Справедливость такого вывода подтверждается кривыми, приведенными на рис. 12[113].

Таковы, в самых общих чертах, взаимосвязи информатики с другими дисциплинами. Эти взаимосвязи показывают, что информатика как самостоятельная дисциплина формируется на стыке многих наук и поэтому широко пользуется их методами. В подобном же положении находятся и другие недавно возникшие

[113] Current research and development in scientific documentation. № 1—11, 13. Washington, National Science Foundation, 1957—1964; American Documentation Institute. Committee on Organisation of Information. Report for 1962—1963. «American Documentation», 1964, v. 15, № 4, стр. 279.

научные дисциплины. Возникновение новой науки на стыке других наук — это сложный и длительный процесс. На это указывал академик А. Н. Несмеянов.

«Начало процесса соприкосновения наук, — писал он, — происходит не так легко. И это понятно, так как подобное взаимодействие требует наличия ученых, которые владеют той или другой из соприкасающихся отраслей знаний, что при нынешней глубокой дифференциации науки встречается не часто. Вот почему важная задачи заключается в подготовке таких специалистов на всех ступенях их квалификации, которые владели хотя бы двумя специальными областями и могли осуществлять перенесение идей и методов одной науки в другую» [114].

---

[114] А. Н. Несмеянов. На стыке наук. «Наука и жизнь», 1958, № 3, стр. 2.

Глава 2

# ДОКУМЕНТАЛЬНЫЕ ИСТОЧНИКИ НАУЧНОЙ ИНФОРМАЦИИ

## НАУЧНЫЕ ДОКУМЕНТЫ И ИХ ВИДЫ

В научно-информационной деятельности важнейшим источником научной информации и средством ее передачи в пространстве и времени служит и будет еще долго служить научный документ. Поэтому важно определить информационное значение основных видов научных документов.

Документы используются в различных областях человеческой деятельности — политической, хозяйственной, научной и являются предметом изучения разных научных дисциплин. Специалисты каждой из этих дисциплин, будь то юрист, искусствовед или архивист, вкладывают в термин «документ» разное содержание.

В широком смысле документами можно считать не только надписи, рукописи и печатные издания, но и произведения искусства, нумизматические памятники, музейные экспонаты минерального, ботанического, зоологического или антропологического характера. *Документом* признается любой материальный объект, который фиксирует или подтверждает какие-либо знания и может быть включен в определенное собрание.

*Научным документом* считается материальный объект, содержащий закрепленную научную информацию, предназначенный для ее передачи во времени и пространстве и используемый в общественной практике. По форме различают документы: текстовые (книги, журналы, рукописи и т. д.), графические, или изобразительные (чертежи, схемы, графики, планы, карты, диаграммы и т. п.), аудиовизуальные (звукозаписи, кинофильмы, диапозитивы и др.).

Типология научных документов разработана недостаточно четко. Различные виды научных документов возникали в разное время и на протяжении последних десятилетий претерпевают значительную эволюцию как по форме, так и по своим функциям. Информатика выдвинула на первый план деление документов на

гораздо более узко и подразумеваем под книгой многостраничное произведение печати. В ряде зарубежных стран книгой признается лишь произведение печати, имеющее объем свыше трех печатных листов, т. е. не менее 49 страниц. Произведения меньшего объема считаются брошюрами. Подобное деление поддерживается и в рекомендациях ЮНЕСКО относительно статистики печати; деление произведений печати на книги и брошюры производится с 1964 г. и в отечественной практике[1].

Книжная статистика — одна из самых малоупорядоченных областей библиографической деятельности. В настоящее время мы не располагаем сколько-нибудь точными цифрами о числе книг, опубликованных со времени изобретения книгопечатания. Приводимые некоторыми авторами цифры — 30, 50 или даже 100 млн. названий книг — не подкреплены доказательствами и нуждаются в разъяснениях. По-видимому, в эти цифры включены в той или иной мере сведения о специальных видах технических изданий, а возможно, и о периодических изданиях. Наиболее достоверными могут считаться данные, полученные из анализа библиографических источников. Известный бельгийский документалист Поль Отле в 1900 г. считал, что со времени изобретения книгопечатания всего было издано свыше 12 млн. названий книг[2].

Польский книговед Болеслав Ивинский в 1911 г., умело применив метод интерполяции для периодов, за которые полные статистические данные отсутствуют (с середины XVI в. до середины XVIII в.), пришел к выводу, что до 1908 г. включительно было опубликовано не более 10,4 млн. названий книг[3]. Последнее по времени капитальное исследование этого вопроса принадлежит американцу Ле Рой Мерриту, который основывался на данных Б. Ивинского. В 1942 г. он оценил мировую книжную продукцию по 1940 г. включительно в 15,4 млн. названий[4]. Если принять эту цифру за исходную и основываться на данных ЮНЕСКО, по которым ежегодно на мировой книжный рынок выпускается 350 тыс. изданий, то можно говорить о 23—25 млн. названий книг, напечатанных со времени Гутенберга.

Однако фактическая цифра должна быть значительно выше, так как число изданий, не поступающих в продажу, в некоторых странах почти в три раза превышает число книг, выпускаемых на

---

[1] Unesco. Report of the Committee of Experts for the International Standardization of Statistics Relating to Book Production and Periodicals. Paris, 1961, стр. 3. Печать СССР в 1966 г. Статистические материалы. М., «Книга», 1967, стр. III.

[2] P. Otlet. La statistique internationale des imprimés. «Bull. de l'Institut international de bibliographie». Bruxelles, 1900, стр. 116 (табл. 4).

[3] B. Iwiński. La statistique internationale des imprimés. «Bull. de l'Institut international de bibliographie». Bruxelles, 1911, стр. 6 (табл. 1).

[4] Le Roy C. Merrit. The administrative, fiscal and quantative aspects of the regional union catalogs in the United States. Chicago, 1942, стр. 72—82 (табл. 29).

книжный рынок. Так, например, в США в 1959 г. из общего числа 60 тыс. названий книг около 45 тыс. названий приходилось на обычно не поступающие в продажу издания университетов, научно-исследовательских и правительственных учреждений, различных обществ и иных организаций [5]. Таким образом, цифры в 30 и даже 50 млн., хотя они и не подкреплены документально, вряд ли можно считать чрезмерно преувеличенными.

Мы уделили данному вопросу особое внимание, так как по поводу оценки суммарного количества опубликованных в мире книг нередко возникают острые дискуссии.

В огромной массе выходящих книг удельный вес книг естественнонаучного содержания относительно скромен — всего 8%. Несколько больше — 13% — издается книг по технике, сельскому хозяйству и медицине. Вместе взятые, они не составляют и половины книг по общественным и гуманитарным дисциплинам (около 50% всех книг) [6]. Эти цифры подтверждаются и на материале отдельных дисциплин. Так, например, при подсчете относительного числа источников по электротехнике выяснилось, что на книги приходится 12% [7]. Место книги среди научных публикаций можно оценивать также по масштабам ее использования учеными и инженерами. Опрос ученых, работающих в американских университетах, и инженеров, занятых в электропромышленности Великобритании, показал, что из всей читаемой ими специальной литературы монографии и учебники в первом случае составляют 19%, а во втором — 14% [8].

Роль книги как источника и средства распространения научных сведений ослабляется тем, что для написания книги даже коллективом авторов требуется несколько лет. Издание ее также занимает длительное время. Поэтому многие сведения, сообщаемые в момент опубликования книги, уже имеют давность от трех до пяти лет и часто оказываются устаревшими. В таких быстро растущих отраслях, как физика и химия, книги стареют быстрее; в них используются главным образом книги последних лет издания. По подсчетам американского библиографа Ч. Брауна, 64% ссылок на книги в работах по этим отраслям науки приходится на издания не старше десяти лет. Определенное представление о темпах старения книг по естественным наукам [9] может быть получено из табл. 2.

---

[5] Council on library resources. 3rd Annual Report for the period ending June 30, 1959. Washington, 1959.

[6] R. E. Barker. Le livre dans le monde. Etude sur le commerce international du livre. Paris, Unesco, 1957.

[7] М. Цвганник. Определение оперативного поискового фонда. «Научно-техн. информация», 1964, № 4, стр. 10—16.

[8] B. C. Vickery. The use of scientific literature. «Library Associations Record», 1961, v. 63, № 8, стр. 263—269.

[9] C. H. Brown. Scientific serials. Chicago, ACRL, 1956, стр. 161.

гораздо более узко и подразумеваем под книгой многостраничное произведение печати. В ряде зарубежных стран книгой признается лишь произведение печати, имеющее объем свыше трех печатных листов, т. е. не менее 49 страниц. Произведения меньшего объема считаются брошюрами. Подобное деление поддерживается и в рекомендациях ЮНЕСКО относительно статистики печати; деление произведений печати на книги и брошюры производится с 1964 г. и в отечественной практике [1].

Книжная статистика — одна из самых малоупорядоченных областей библиографической деятельности. В настоящее время мы не располагаем сколько-нибудь точными цифрами о числе книг, опубликованных со времени изобретения книгопечатания. Приводимые некоторыми авторами цифры — 30, 50 или даже 100 млн. названий книг — не подкреплены доказательствами и нуждаются в разъяснениях. По-видимому, в эти цифры включены в той или иной мере сведения о специальных видах технических изданий, а возможно, и о периодических изданиях. Наиболее достоверными могут считаться данные, полученные из анализа библиографических источников. Известный бельгийский документалист Поль Отле в 1900 г. считал, что со времени изобретения книгопечатания всего было издано свыше 12 млн. названий книг [2].

Польский книговед Болеслав Ивинский в 1911 г., умело применив метод интерполяции для периодов, за которые полные статистические данные отсутствуют (с середины XVI в. до середины XVIII в.), пришел к выводу, что до 1908 г. включительно было опубликовано не более 10,4 млн. названий книг [3]. Последнее по времени капитальное исследование этого вопроса принадлежит американцу Ле Рой Мерриту, который основывался на данных Б. Ивинского. В 1942 г. он оценил мировую книжную продукцию по 1940 г. включительно в 15,4 млн. названий [4]. Если принять эту цифру за исходную и основываться на данных ЮНЕСКО, по которым ежегодно на мировой книжный рынок выпускается 350 тыс. изданий, то можно говорить о 23—25 млн. названий книг, напечатанных со времени Гутенберга.

Однако фактическая цифра должна быть значительно выше, так как число изданий, не поступающих в продажу, в некоторых странах почти в три раза превышает число книг, выпускаемых на

---

[1] Unesco. Report of the Committee of Experts for the International Standardization of Statistics Relating to Book Production and Periodicals. Paris, 1961, стр. 3. Печать СССР в 1966 г. Статистические материалы. М., «Книга», 1967, стр. III.

[2] P. Otlet. La statistique internationale des imprimés. «Bull. de l'Institut international de bibliographie». Bruxelles, 1900, стр. 116 (табл. 4).

[3] B. Iwiński. La statistique internationale des imprimés. «Bull. de l'Institut international de bibliographie». Bruxelles, 1911, стр. 6 (табл. 1).

[4] Le Roy C. Merrit. The administrative, fiscal and quantative aspects of the regional union catalogs in the United States. Chicago, 1942, стр. 72—82 (табл. 29).

книжный рынок. Так, например, в США в 1959 г. из общего числа 60 тыс. названий книг около 45 тыс. названий приходилось на обычно не поступающие в продажу издания университетов, научно-исследовательских и правительственных учреждений, различных обществ и иных организаций [5]. Таким образом, цифры в 30 и даже 50 млн., хотя они и не подкреплены документально, вряд ли можно считать чрезмерно преувеличенными.

Мы уделили данному вопросу особое внимание, так как по поводу оценки суммарного количества опубликованных в мире книг нередко возникают острые дискуссии.

В огромной массе выходящих книг удельный вес книг естественнонаучного содержания относительно скромен — всего 8%. Несколько больше — 13% — издается книг по технике, сельскому хозяйству и медицине. Вместе взятые, они не составляют и половины книг по общественным и гуманитарным дисциплинам (около 50% всех книг) [6]. Эти цифры подтверждаются и на материале отдельных дисциплин. Так, например, при подсчете относительного числа источников по электротехнике выяснилось, что на книги приходится 12% [7]. Место книги среди научных публикаций можно оценивать также по масштабам ее использования учеными и инженерами. Опрос ученых, работающих в американских университетах, и инженеров, занятых в электропромышленности Великобритании, показал, что из всей читаемой ими специальной литературы монографии и учебники в первом случае составляют 19%, а во втором — 14% [8].

Роль книги как источника и средства распространения научных сведений ослабляется тем, что для написания книги даже коллективом авторов требуется несколько лет. Издание ее также занимает длительное время. Поэтому многие сведения, сообщаемые в момент опубликования книги, уже имеют давность от трех до пяти лет и часто оказываются устаревшими. В таких быстро растущих отраслях, как физика и химия, книги стареют быстрее; в них используются главным образом книги последних лет издания. По подсчетам американского библиографа Ч. Брауна, 64% ссылок на книги в работах по этим отраслям науки приходится на издания не старше десяти лет. Определенное представление о темпах старения книг по естественным наукам [9] может быть получено из табл. 2.

---

[5] Council on library resources. 3rd Annual Report for the period ending June 30, 1959. Washington, 1959.

[6] R. E. Barker. Le livre dans le monde. Etude sur le commerce international du livre. Paris, Unesco, 1957.

[7] М. Цыганик. Определение оперативного поискового фонда. «Научно-техн. информация», 1964, № 4, стр. 10—16.

[8] B. C. Vickery. The use of scientific literature. «Library Associations Record», 1961, v. 63, № 8, стр. 263—269.

[9] C. H. Brown. Scientific serials. Chicago, ACRL, 1956, стр. 161.

ства. Например, «Распорядок работы учреждений Академии наук СССР» — ведомственное издание, так как оно представляет интерес лишь для сотрудников учреждений АН СССР. С другой стороны, «План работы научно-исследовательских институтов АН СССР» — официальное издание более широкого назначения, поскольку оно может интересовать научных работников, независимо от их ведомственного подчинения.

Учреждения и организации публикуют и литературу, которая не носит характера официальных изданий. Монографии, например, сборники статей и докладов, учебники и руководства, издаваемые различными научными учреждениями, не являются их официальными изданиями, хотя и публикуются от их имени, так как по своему содержанию они не связаны с их внутренней деятельностью и могут отражать личные точки зрения авторов на излагаемые вопросы. Многие официальные издания не представляют научного интереса. Однако часть из них содержит важную научную или техническую информацию. Это — отчеты и планы, бюджеты, обзоры работы или докладные записки о работе научных и других специальных учреждений или организаций, их постановления, решения и резолюции, ведомственные нормативные материалы (часть из них рассматривается как специальные виды технических изданий), производственные и технологические инструкции, информационно-справочные, методические и статистические материалы, программы и учебные планы и т. п.

Как уже было сказано, все перечисленные виды книг могут быть не только однотомными, но и многотомными изданиями, а также составлять серии. Для практической научно-информационной работы важно отличать многотомные издания от серийных. Многотомным считается издание, опубликованное в двух или более томах, книгах, частях или выпусках, составляющих — в отличие от серии — единое целое, ограниченное определенным числом томов. *Серия* состоит из ряда самостоятельных произведений печати, обычно связанных общностью или близостью тематики и выпускаемых в свет одним издательством под общим серийным заглавием. В отличие от многотомников, тома которых обычно нумеруются, выпуски серии могут быть ненумерованными. За рубежом многие научные издания, как выходящие в коммерческих издательствах, так и выпускаемые научными учреждениями и организациями, объединяются в серии [12]. Часто серии распадаются на подсерии. В отечественной книгоиздательской практике серийными чаще являются научно-популярные издания, рассчитанные не на специалистов, а на широкие круги читателей.

---

[12] При работе с зарубежными изданиями на английском языке следует учитывать, что термин «serials» не совпадает с русским понятием «серия» и служит для обозначения периодических изданий.

## Периодические издания

*Периодическим изданием* принято считать произведение печати, выходящее регулярно через определенные или неопределенные промежутки времени отдельными, не повторяющимися по содержанию выпусками под единым объединяющим их заглавием. Периодическое издание, как правило, рассчитано на заранее неограниченную продолжительность выхода в свет и обычно имеет одинаковое оформление. Традиционными видами периодических изданий являются журналы и газеты. К периодическим изданиям относят также продолжающиеся издания, которые занимают промежуточное положение между книгами и журналами.

*Продолжающиеся издания* — это сборники научных работ и других материалов учреждений, обществ или иных организаций, рассчитанные на постоянное или длительное существование, публикуемые без строгой периодичности нумерованными выпусками под одним общим заглавием («Труды», «Записки», «Известия» и т. п.) и имеющие, как правило, единообразное оформление. На практике некоторые продолжающиеся издания трудно отличать, с одной стороны, от книжных серий, а с другой — от журналов. Важным признаком продолжающихся изданий, позволяющим выделить их из книжных серий, служит указание на то, что они выпущены учреждениями, которые не являются издательствами. Чаще всего продолжающиеся издания выпускаются академиями, университетами, научно-исследовательскими и учебными институтами, научными обществами и т. п. Другой признак, который отличает продолжающиеся издания от книг-многотомников, состоит в том, что такие издания не имеют законченного плана и заранее установленного числа выпусков (томов).

Помимо изданий, выходящих без строгой периодичности, к продолжающимся относятся также издания, выходящие через равные промежутки времени, но не чаще, чем один раз в год (ежегодники и издания, выходящие один раз в два, три, четыре, пять и более лет). Этим они отличаются от журналов, которые в течение года выходят через строго установленные промежутки времени или же имеют заранее установленное число номеров в год. Продолжающиеся издания — важный источник научной информации: в них часто сообщаются сведения, не проникающие в широко распространенные журналы, публикуются доклады, статьи и материалы по узкоспециальным вопросам, отражающим направление исследований того или иного учреждения.

*Журналом* называется периодическое издание, регулярно выходящее еженедельными, ежемесячными, двухмесячными, квартальными или полугодовыми выпусками, одинаково оформленными и содержащими статьи или иные материалы по научно-техническим, общественно-политическим, художественным и другим вопросам. В журналах содержится новейшая информация,

освещающая последние достижения в науке, а также новые методы и изобретения в технике. Научный журнал — старейший вид периодических изданий. Его начало восходит к французскому «Journal des sçavants», первый номер которого вышел 5 января 1665 г. В течение второй половины XVII в. и всего XVIII в. научный журнал являлся главным образом источником информации о новых научных книгах. Вначале оригинальные статьи публиковались в журналах сравнительно редко. Обычно они принадлежали крупным ученым и носили традиционно-условную форму писем одного ученого к другому (так было принято сообщать о научных открытиях в предшествующие эпохи). Начиная с XIX в., журнал становится основным источником информации о новых достижениях в науке и технике. В связи с этим быстро растет число названий научно-технических журналов.

Среди источников, которыми пользуются ученые и инженеры, журнальные статьи прочно занимают первое место. Обследование библиографических запросов нескольких тысяч американских ученых показало, что 68% этих запросов приходится на журнальные статьи. Из источников, которыми пользуются инженеры электропромышленности Великобритании, 70% падает на специальные журналы; 60% всех научных документов по электротехнике составляют журнальные статьи [13].

«Свод рекомендуемых правил в области научных публикаций», принятый в 1961 г. Международной организацией по стандартизации и одобренный ЮНЕСКО, делит все статьи в журналах по естественным, точным и техническим наукам на три категории [14]:

а) оригинальные научные труды (*mémoires scientifiques originaux*), которые позволяют научным работникам, не обращаясь к другим источникам, воспроизвести поставленные автором опыты, повторить его наблюдения, проверить точность его наблюдений и выводов;

б) временные публикации и предварительные замечания (publications provisoires ou notes initiales), которые могут содержать элементы научной новизны, но не отвечают условиям, сформулированным в пункте «а»;

в) уточняющие заметки (exposes de mise au point), которые не содержат новой научной информации; в них приводятся, анализируются и обсуждаются опубликованные ранее сведения [15].

Рост числа публикаций в научных журналах, недостаточная профилированность этих журналов, быстрое старение опубликованных материалов, их неудовлетворительный библиографический

[13] R. S h a w. Pilot study on the use of scientific literature by scientists. Washington, 1956, стр. 15.

[14] Свод рекомендуемых правил в области научных публикаций. «Бюлл. ЮНЕСКО для библиотек», 1963, т. 17, стр. 26—32.

[15] Code du bon usage en matiere de publications scientifiques. «Bull. de l'Unesco a l'intention des bibliotheques», 1963, v. 17, № 1, стр. 30—34.

учет, аннотирование и реферирование привели к тому, что журнал как источник научной информации подвергается все большей критике.

Еще в 1958 г. на Конференции по научной информации, созванной Лондонским королевским обществом, известный английский физик Дж. Бернал предложил план преобразования системы публикаций в научных журналах, который сводился к следующему. Вместо многочисленных журналов он предложил создать ограниченное число специализированных информационно-реферативных центров сначала в национальном, а затем и в международном масштабе. Все статьи должны посылаться в эти центры. Там эти статьи редактируются, реферируются и передаются на хранение. Рефераты статей публикуются в реферативном журнале. Полные тексты статей могут быть высланы читателю по его требованию. Все процессы работы в таком информационном центре, вплоть до реферирования, должны быть со временем механизированы [16].

«План Бернала», естественно, вызвал много различных споров. Ряд зарубежных ученых, документалистов и библиотекарей поддерживает предложение Дж. Бернала и разрабатывает детали реализации этого плана. Однако сам Дж. Бернал пошел еще дальше. На Международной конференции по научной информации, состоявшейся в Вашингтоне в 1958 г., он предложил отказаться и от статьи как формы сообщения результатов исследования. Дж. Бернал считает, что многие статьи содержат факты и выводы, полученные в результате исследований, которые проводились в течение длительного времени. По его мнению, информационный центр для ускорения передачи научной информации мог бы собирать не статьи, а непосредственно результаты и выводы каждого исследования [17]. Едва ли нужно доказывать, что практическая реализация «плана Бернала» не только крайне затруднительна, но и не во всем целесообразна. Однако этот план содержит отдельные рациональные моменты (например, предложение о депонировании некоторых видов научно-технической литературы).

Известный интерес в этом отношении представляет решение проблемы создания централизованных депозитных хранилищ, которое было предложено Американским физическим обществом. Все научные труды, поступающие в это общество, подразделяются на две категории: а) актуальные и в соответствии с этим представляющие большую ценность и б) менее актуальные и поэтому не требующие быстрого опубликования. Первые издаются

[16] J. D. Bernal. Provisional scheme for central distribution of scientific publications. В кн.: «The Royal Society Scientific Information Conference». London, 21 June — 2 July 1948. Reports and papers submitted. London, 1948.

[17] J. D. Bernal. Scientific informtion and its users. «ASLIB Proceedings», 1960, v. 12, № 12, стр. 432—438.

в сокращенном виде менее чем через месяц в бюллетене «Physical Review Letters». Вторые печатаются полностью через четыре—шесть месяцев в журнале «Physical Review», причем рефераты помещенных в этом журнале статей публикуются в приложении к бюллетеню «Physical Review Letters». Журнал «Physical Review» становится своего рода центральным депозитным хранилищем статей [18].

Известно, что одной из основных задач информационных служб является обеспечение исчерпывающей полноты охвата научно-технической литературы по тому или иному вопросу. Для этого каждая информационная служба стремится приобрести все издания по интересующему ее вопросу, причем решения о приобретении периодических изданий принимаются главным образом на основании их названий и тематики. Однако такие критерии в принципе не позволяют обеспечивать исчерпывающей полноты сбора публикаций по любому вопросу.

Еще в начале 30-х годов нашего века английский химик и библиограф С. Брэдфорд обнаружил, что примерно 50% всех статей по электротехнике публикуется в небольшом числе (около 200) журналов, которые являются «профильными» для данной отрасли науки. Кроме того, не менее 25% статей по электротехнике помещается в значительно большем числе периодических изданий, непосредственно не относящихся к электротехнике. К числу таких изданий относятся журналы по физике, энергетике, транспорту и многим другим отраслям науки и техники. Остальные статьи по электротехнике публикуются в большом числе периодических изданий, вообще не имеющих никакого отношения к электротехнике.

Появление публикации по какой-либо отрасли науки или техники в журналах, не имеющих прямого или косвенного отношения к этой отрасли, объясняется единством знаний. Каждая отрасль науки в большей или меньшей степени связана с любой другой отраслью науки. Поэтому время от времени в периодических изданиях, посвященных какой-либо отрасли науки или техники, могут публиковаться статьи, представляющие интерес для специалистов других отраслей. Очевидно, что число таких периодических изданий будет расти по мере ослабления связей, которые существуют между интересующей нас отраслью науки или техники и другими отраслями. Наоборот, число публикаций по данной отрасли, помещаемых в таких периодических изданиях, будет соответственно уменьшаться.

Это означает, что все периодические издания можно распределить по концентрическим зонам уменьшающейся продуктивности

---

[18] Science, government, and information. A report of the President's Science Advisory Committee. Washington, U. S. Government Printing Office, 1963, стр. 31—32.

(в отношении публикаций по некоторой данной отрасли науки или техники). Тогда число периодических изданий в каждой последующей зоне будет увеличиваться, а их продуктивность соответственно уменьшаться.

На основании изложенных выше фактов и соображений С. Брэдфорд выдвинул следующую гипотезу. Пусть имеется некоторое число журналов, посвященных непосредственно данной отрасли или предмету. Эти журналы образуют некоторое ядро (nucleus). Вокруг такого ядра можно построить кольцевые зоны, каждая из которых содержит столько же публикаций по данному предмету, что и ядро. В первую зону попадут журналы, которые лишь частично посвящены рассматриваемой отрасли или предмету. Таких журналов будет насчитываться в $n$ раз больше, чем журналов в ядре. Вторую зону образуют журналы, еще в меньшей степени относящиеся к данной отрасли или предмету, чем журналы первой зоны. Они будут находиться примерно в таком же отношении к тематике журналов первой зоны, в каком журналы первой зоны находятся к тематике журналов ядра. Поэтому можно предположить, что во второй зоне содержится в $n^2$ больше журналов, чем в ядре, и т. д.

Таким образом С. Брэдфорд получил модель, которая достаточно хорошо описывала линейный характер рассеяния публикаций в периодических изданиях. Эта модель имеет следующий вид:

$$p : p_1 : p_2 = 1 : n : n^2,$$

где $p$, $p_1$ и $p_2$ — число журналов в ядре и последующих зонах и

$$n = \frac{p_1}{p} = \frac{p_2}{p_1} \approx 5.$$

С. Брэдфорд дал следующую формулировку открытому им закону рассеяния публикаций: «Если научные журналы расположить в порядке уменьшения числа помещенных в них статей по какому-либо заданному предмету, то в полученном списке можно выделить ядро журналов, посвященных непосредственно этому предмету, и несколько групп или зон, каждая из которых содержит столько же статей, что и ядро. Тогда числа журналов в ядре и последующих зонах будут относиться как $1 : n : n^2$» [19].

Для проверки своего закона С. Брэдфорд исследовал характер фактического рассеяния 1332 статей по прикладной геофизике в 326 журналах за 1928—1931 гг., а также 395 статей по смазке в 164 журналах за 1931—1933 гг. Исследование показало, что если расположить журналы в порядке уменьшения их продуктивности в отношении статей по какому-либо предмету, то суммарное число помещенных в них статей будет пропорционально суммар-

[19] S. C. Bradford. Documentation, 2nd ed. London, Lockwood, 1953, стр. 154.

ному числу этих журналов (или его логарифму). Эта зависимость не распространяется лишь на журналы, образующие ядро.

Однако в 1948 г. Б. Викери показал, что при использовании модели Брэдфорда график зависимости $\Sigma p_i a_i$ от $\Sigma p_i$ (или $\lg \Sigma p_i$), где $p_i$ — число журналов $i$-того ранга, а $a_i$ — число статей по данному предмету в журналах $i$-того ранга, имеет небольшую кривизну [20]. Для устранения этой кривизны Б. Викери внес уточнение в модель Брэдфорда, которая после этого получила следующий вид:

$$T_x : T_{2x} : T_{3x} = 1 : n : n^2,$$

где $T_x$ — число журналов, содержащее $x$ статей по данному предмету; $T_{2x}$ — общее число журналов, содержащих $2x$ таких статей; $T_{3x}$ — общее число журналов, содержащих $3x$ статей; $n$ — любое число, зависящее от величины, которая выбрана для $x$.

В 1960 г. английский математик М. Кендалл проанализировал рассеяние 1763 статей по исследованию операций в 370 журналах. Результаты этого исследования показали, что первые 5 наиболее продуктивных журналов содержали почти 35% всех статей по исследованию операций, помещенных во всех этих журналах, 18 журналов — 50% и 67 журналов — 75% всех статей по исследованию операций. Построенный М. Кендаллом график зависимости $\Sigma p_i a_i = f(\lg \Sigma p_i)$ отличается даже лучшей линейностью, чем графики С. Брэдфорда. Аналогичная картина рассеяния 1465 публикаций по методологии статистики была выявлена М. Кендаллом в 164 журналах за 1925—1939 гг.

Вместе с тем М. Кендалл показал, что рассеяние брэдфордовского типа в сущности является лишь частным случаем распределения, описываемого так называемым законом Ципфа [21].

Для аналитического выражения закона Ципфа предположим, что к некоторому тексту составлен словник объемом в $L$ слов, причем для каждого слова указана частота его встречаемости в данном тексте. Расположим эти слова в порядке убывания частот и пронумеруем их от 1 (номер самого частого слова) до $V$. Номер каждого слова в этом списке будем называть его рангом. Тогда между частотой встречаемости слов в тексте и рангами этих слов устанавливается следующая зависимость:

$$p_r = \frac{0{,}1}{r},$$

[20] B. C. Vickery. Bradford's law of scattering. «Journal of Documentation», 1948, v. 4, № 3, стр. 198—203.

[21] M. G. Kendall. The bibliography of operational research. «Operational Research Quarterly», 1960, v. 11, № 1—2, стр. 31—36; F. F. Leimkuhler. The Bradford distribution. «Journal of Documentation», 1967, v. 23, № 3, стр. 197—207.

где $p_r$ — вероятность появления $r$-того слова в тексте и $r$ — ранг этого слова.

Закон Ципфа выражает некоторое универсальное свойство, которое присуще всем естественным языкам мира. Это свойство заключается в том, что основную часть любого текста на любом языке образует весьма небольшое число часто употребляемых слов. Остальную часть текстов составляют десятки тысяч слов, которые употребляются очень редко. Например, всего 300 слов образуют около 65% всех текстов, 500 слов — около 70%, 1000 слов — 80%, 2000 слов — 86% и т. д. Отсюда следует, что для построения 100% всех текстов необходимо иметь словарь, в котором насчитываются сотни тысяч редко употребляемых слов языка.

Исследования показали, что закон Ципфа применим не только к словам, но и к более длинным сегментам текстов — к фразам, статьям и даже журналам. В применении к журналам эта зависимость нашла свое выражение в законе рассеяния Брэдфорда, о котором говорилось выше. Этот вывод был подтвержден также результатами изучения характера рассеяния публикаций по физике и химии (с химической технологией), которое было проведено авторами в конце 1966 г.

Явление рассеяния публикаций исследовали многие специалисты [22]. В сводном виде результаты основных из этих исследований приведены в табл. 3.

Рассмотрение приведенных в табл. 3 данных показывает, что наименьшее рассеяние публикаций наблюдается в теоретических науках. Для отраслей техники и производства оно увеличивается и достигает максимума для общественных наук.

Закон рассеяния публикаций имеет исключительно большое практическое значение. Из этого закона прежде всего следует, что охват всех публикаций по какой-либо отрасли или предмету не может быть обеспечен, если ограничиться просмотром лишь профильных журналов и журналов по родственной тематике. Для обеспечения исчерпывающей полноты реферирования публикаций по той или иной отрасли или предмету информационная служба должна просматривать практически все научно-техниче-

---

[22] F. F. Leimkuhler. The Bradford distribution. «Journal of Documentation», 1967, v. 23, № 3, стр. 198; S. C. Bradford. Documentation. 2nd ed., London, Lockwood, 1953, стр. 150—151; M. G. Kendall. The bibliography of operational research. «Operational Research Quarterly», 1960, v. 11, № 1—2, стр. 31—36; P. M. Bernays, K. L. Coe and J. L. Wood. A computer-based source inventory of «Chemical Abstracts». «Journal of Chemical Documentation», 1965, v. 5, № 4, стр. 242—249. Chemical literature expands. «Chemical and Engineering News», 1966, v. 44, № 23, стр. 88; P. F. Cole. A new look at reference scattering. «Journal of Documentation», 1962, v. 18, № 2, стр. 58—64; S. Keenan and P. Atherton. The journal literature of physics. A comprehensive study based on «Physics Abstracts» (Science Abstracts, Section A) 1961 issue AIP/DRP PAI (1964). New York, The American Institute of Physics, 1964.

Таблица 3

**Рассеяние публикаций по различным отраслям науки, техники и производства**

| Отрасль науки, техники или производства | Число исследованных публикаций | Число периодических изданий, содержащих публикации | | | |
|---|---|---|---|---|---|
| | | 25% | 50% | 75% | 100% |
| Химия | 3633 | 2 | 7 | 24 | 247 |
| | 1085 | 1 | 5 | 19 | 131 |
| Химия и химическая технология | 187911 | 50 | 250 | 1000 | 10850 |
| Химическая технология | 21728 | 3 | 23 | — | — |
| Биохимия | 17198 | 3 | 12 | 56 | 851 |
| Физика | 1279 | 1 | 3 | 17 | 134 |
| | 20287 | 7 | 19 | 47 | 405 |
| Прикладная геофизика | 1332 | 5 | 27 | 85 | 326 |
| Исследование операций | 1763 | 3 | 18 | 67 | 370 |
| Методология статистики | 1465 | 2 | 7 | 18 | 164 |
| Электротехника | 17991 | 3 | 9 | 39 | — |
| Радиотехника | 1506 | 2 | 8 | 20 | — |
| Нефтяная промышленность | 903 | 5 | 15 | 69 | 197 |
| Смазка | 164 | 7 | 22 | 62 | 395 |
| История США | 452 | 14 | 54 | 149 | 259 |

ские журналы, число которых в настоящее время превысило 30 тыс. названий. Но это возможно лишь при условии централизованной обработки научно-технической литературы, которое характерно для нашей страны.

Кроме того, закон Брэдфорда позволяет определить число периодических изданий, которые обеспечивают тот или иной процент всех публикаций по какой-либо отрасли или предмету. Это особенно важно при организации выпуска изданий сигнальной информации. Например, 18 наиболее продуктивных журналов обеспечивают 15% всех публикаций по химии и химической технологии, 50 журналов — 25%, 100 журналов — 34%, 500 журналов — 62%, 1000 журналов — 75% и т. д.

Большую ценность закон Брэдфорда представляет для библиотек. Дело в том, что этот закон распространяется не только на журналы, но и на книги и другие виды научно-технических изданий. Это позволяет комплектовать фонды библиотек в расчете на разные степени удовлетворения потребностей читателей. Например, обследования показали, что для удовлетворения 90—95% запросов читателей библиотека медицинского училища должна иметь фонд около 70—90 тыс. томов, тогда как для удовлетворе-

ния 100% запросов этих же читателей потребовался бы фонд в 0,4—0,5 млн. томов [23].

Наконец, изучение характера рассеяния публикаций и особенно — его изменений во времени позволяет выявлять различные тенденции в развитии исследований в той или иной отрасли науки или техники.

Интересно проследить за старением журнальных статей. По сравнению с книгами, они публикуются значительно быстрее. Однако статьи содержат меньше обобщений и больше конкретных фактов и потому устаревают даже несколько скорее книг. Для характеристики процесса старения журнальных статей в определенной отрасли науки недавно было применено понятие, заимствованное из ядерной физики. По аналогии с периодом полураспада радиоактивных веществ стали исчислять полупериоды жизни научной литературы.

Полупериодом считают время, в течение которого была опубликована половина всей используемой в настоящий момент литературы в определенной области. Разумеется, используемость литературы определяется условно, по косвенным показателям, например, по читательским запросам в библиотеке или чаще всего по библиографическим ссылкам. Полупериод жизни статей в научных журналах точно совпадает со временем, в течение которого половина всей опубликованной в настоящий момент литературы в определенной области перестанет использоваться. Это обстоятельство имеет очень важное значение, так как на основе анализа старения литературы по любой дисциплине или вопросу в прошлом можно делать достоверные прогнозы относительно сравнительной ценности публикуемой литературы в будущем.

Ориентировочные подсчеты полупериодов жизни статей в различных отраслях науки дают следующие результаты (в годах) [24]:

| | | | |
|---|---|---|---|
| Физика | — 4,6 | Ботаника | — 10,0 |
| Физиология | — 7,2 | Математика | — 10,5 |
| Химия | — 8,1 | Геология | — 11,8 |

Таким образом, значительная часть журнальных статей по физике стареет очень быстро, является как бы эфемерной литературой, в физиологии и химии быстро устаревающих статей примерно столько же, сколько и долговечных, а в ботанике, математике и геологии преобладают долговечные работы.

Имеется возможность вывести и общие средние данные. На рис. 13 показано распределение работ, на которые были сделаны ссылки в 296 тыс. зарубежных научных публикациях 1965 г.,

[23] Esterguest report proposes. New York State aid to medical libraries. «Bull. of the Medical Libraries Association», 1963, v. 51, № 1, стр. 125—127.

[24] R. E. Burton and R. E. Kebler. The «half-life» of some scientific and technical literatures. «American Documentation», 1960, v. 11, № 1, стр. 20.

по точным, естественным и прикладным наукам. Получается следующая картина распределения цитированных работ по их возрасту. Наибольшее число упоминаемых работ (13% общего количества) относится к 1964 г. Опубликованные работы не сразу становятся известны исследователям и потому публикации 1965 г.

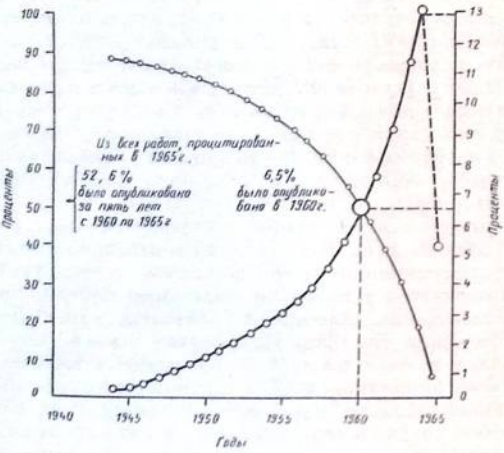



Рис. 13. Распределение по годам опубликования работ, цитированных в журнальных статьях 1965 г.

упоминаются значительно реже (5% общего числа работ, на которые ссылались в 1965 г.). Затем следует спад, середина которого (6,5% работ) приходится на 1960 г. Работы, опубликованные с 1960 по 1965 г., составляют 52,6% всех работ, на которые в 1965 г. были сделаны ссылки. Следовательно, средний возраст научных работ, составляющих активный исследовательский фронт, не превышает пяти лет.

*Газетами* считаются периодические издания, преимущественно ежедневные, выходящие два или три раза в неделю, еженедельные, в которых публикуются материалы о текущих, главным образом общественно-политических событиях. Как правило, в газетах не публикуется информация научного или технического характера; в них помещаются лишь краткие научно-технические сообщения хроникального типа. Свежесть и новизна сведений, публикуемых в газетах, делают их ценным источником информации. Хорошо налаженная служба газетных вырезок позволяет получать информацию по многим важным вопросам, помещаемую в любой газете страны. Длительные сроки публикации новых материалов

в журналах побуждают некоторых западных ученых пользоваться газетами для сообщений о новых открытиях или достижениях. Однако рекламный стиль, присущий многим газетам буржуазных стран, делает их непригодными для целей научной информации. Неслучайно редактор американского журнала «Physical Review Letters» С. Гоудсмит вынужден был предупредить авторов, что работы, основное содержание которых уже было опубликовано в ежедневной прессе, не могут быть опубликованы в журнале. При этом он заметил, что «научные открытия — предмет, мало подходящий для газетных сенсаций» и что «подобные средства массового распространения служат препятствием для подлинно научной информации» [25].

### Специальные виды технических изданий

Специальные виды технических изданий включают государственные стандарты, межотраслевые и отраслевые технические документы, описания изобретений к авторским свидетельствам и патентам, технические каталоги и прейскуранты на материалы и оборудование, материалы технической информации и обмена опытом. Применяется и другой, более широкий термин «специальные виды технической литературы», который обозначает также и неизданные документы, например, отчеты о научно-исследовательских и опытно-конструкторских работах и переводы статей из иностранных изданий. Заслуга введения и обоснования этого термина принадлежит советскому библиотековеду и библиографу А. Н. Барабанову, который в середине 30-х годов впервые выявил общие черты, объединяющие эти издания по их содержанию, оформлению, происхождению и назначению [26]. В настоящее время оба термина подвергаются критике. Библиограф Д. Ю. Теплов, например, считает, что некоторые из специальных видов технических изданий, в частности технические каталоги и материалы технической информации, являются книгами или периодическими изданиями и не имеют принципиальных от них отличий [27].

*Стандартами* называются научно обоснованные технические документы, которые определяют типы, виды, марки продукции, нормативы ее качества, методы испытаний, упаковку, маркировку, транспортирование и хранение, а также устанавливают общетехнические величины, единицы измерений, термины и обозначения. Обычно стандарты всесторонне характеризуют ту или

---

[25] S. Goudsmit. Editorial. «Physical Review Letters», 1960, № 4, стр. 1—2.

[26] А. Н. Барабанов. Библиотечная работа со специальными видами технической литературы. М., 1936.

[27] Д. Ю. Теплов. Об определении так называемых «специальных видов» технической литературы. «Технические библиотеки СССР», 1963, вып. 2, стр. 23—25.

яную продукцию, но могут касаться и отдельных вопросов, например, методов испытаний.

Стандарты устанавливаются для того, чтобы обеспечить высокое качество промышленной и сельскохозяйственной продукции, для унификации деталей и узлов машин и отдельных изделий. Они позволяют улучшить использование материалов, энергии и топлива, сократить сроки проектирования и снизить себестоимость продукции, служат средством информации о технических достижениях.

В СССР стандарты утверждаются Государственным комитетом стандартов, мер и измерительных приборов и издаются в виде отдельных листов, брошюр или сборников. Отдельно изданные стандарты объемом до одного печатного листа не имеют титульного листа и обложки. Выходные сведения в этом случае располагаются в специальной рамке на первой странице: в левом верхнем углу дается название утвердившей стандарт организации; в середине — наименование вида издания, а под ним заглавие стандарта; в правом верхнем углу — индекс стандарта, состоящий из его сокращенного наименования и номера, например, ГОСТ 8002—62 (первые четыре цифры обозначают регистрационный номер ГОСТ'а, а последние две после тире — год его утверждения). Помимо этих сведений, указывается индекс того стандарта, взамен которого введен настоящий его шифр по «Классификатору общесоюзных стандартов», наименование организации, разработавшей стандарт, даты утверждения и срока введения стандарта, а также его цена

К стандартам близки *межотраслевые и отраслевые технические документы* (технические условия, нормали и другие нормативные документы). В отличие от стандартов, они имеют более ограниченную сферу действия и создаются тогда, когда отсутствуют стандарты на данную продукцию или же когда в силу особых условий производства и эксплуатации возникает необходимость установления специальных требований к этой продукции. С 1960 г. установлены следующие виды технических условий: межреспубликанские (МРТУ), республиканские (РТУ), отдельных предприятий или их групп (ТУ). Нормали устанавливаются на узлы и детали машин, механизмов и приборов, а также на технологическую оснастку. Они определяют конструкцию, размеры, материалы и другие технические данные, необходимые для изготовления деталей и узлов машин. В настоящее время установлены следующие виды нормалей: межотраслевые (МН), отраслевые (ОН) и заводские (Н). К технической документации относятся и руководящие технические материалы (РТМ), выпускаемые отдельными учреждениями и организациями. В отличие от технических условий и нормалей их применение не является обязательным. В них излагаются рекомендуемые методы расчетов, сборки и сварки, способы применения защитных покрытий и т. п.

За рубежом, наряду со стандартами, широко издаются *спецификации*, представляющие собой коммерческие документы, содержащие описания технических требований к продуктам, материалам, процессам или видам обслуживания. Спецификации устанавливают требования, предъявляемые покупателями и потребителями к изготовителям. Зарубежные стандарты, как и советские, различаются по индексам, которые обычно состоят из сокращенного названия выпускающей стандарт организации, номера серии и года утверждения стандарта. Индекс «ASA. B 18,5 : 1952», например, означает: «ASA» — «American Standards Association», «B» — «Машиностроение», 18 — «Болты и гайки», 5 — серийный номер, 1952 — год утверждения стандарта. В капиталистических странах стандарты и примыкающая к ним нормативная литература издаются правительственными и частными организациями и, как правило, имеют лишь рекомендательный характер. В настоящее время стандарты играют важную роль в развитии науки, техники и торговли; они широко используются учеными, исследователями, инженерами, строителями и другими специалистами.

*Описания изобретений* к авторским свидетельствам и патентам (или иначе «патентная литература») являются одним из важных документальных источников научно-технической информации. Они позволяют проследить историю любого изобретения или открытия, а также выяснить современное состояние и направление развития технической мысли в соответствующей области. Современное ознакомление с патентной литературой предотвращает возможность повторного изобретения или разработки уже известных машин и методов. Необходимо указать, что сами авторские свидетельства и патенты являются лишь юридическими документами и не содержат никакой научно-технической информации. Они хранятся в одном экземпляре у изобретателя или патентообладателя и удостоверяют его права на сделанное изобретение.

В СССР основной формой охраны авторских прав изобретателя является авторское свидетельство. Оно обеспечивает бессрочное сохранение авторского права изобретателя на изобретение, а также право на получение вознаграждения и пользование всеми льготами, предоставляемыми в нашей стране изобретателям и рационализаторам. В соответствии с природой социалистического строя, при котором средствами производства владеет государство, авторское свидетельство предоставляет государству исключительное право на использование изобретения.

В капиталистических странах выдаются лишь патенты, предоставляющие их владельцам право в течение 15—17 лет использовать данное изобретение. В СССР патенты обычно выдаются изобретателям, являющимся гражданами иностранных государств.

Как уже было сказано, патентная литература — это не сами авторские свидетельства или патенты, а описания изобретений

иную продукцию, но могут касаться и отдельных вопросов, например, методов испытаний.

Стандарты устанавливаются для того, чтобы обеспечить высокое качество промышленной и сельскохозяйственной продукции, для унификации деталей и узлов машин и отдельных изделий. Они позволяют улучшить использование материалов, энергии и топлива, сократить сроки проектирования и снизить себестоимость продукции, служат средством информации о технических достижениях.

В СССР стандарты утверждаются Государственным комитетом стандартов, мер и измерительных приборов и издаются в виде отдельных листов, брошюр или сборников. Отдельно изданные стандарты объемом до одного печатного листа не имеют титульного листа и обложки. Выходные сведения в этом случае располагаются в специальной рамке на первой странице: в левом верхнем углу дается название утвердившей стандарт организации; в середине — наименование вида издания, а под ним заглавие стандарта; в правом верхнем углу — индекс стандарта, состоящий из его сокращенного наименования и номера, например, ГОСТ 8002—62 (первые четыре цифры обозначают регистрационный номер ГОСТ'а, а последние две после тире — год его утверждения). Помимо этих сведений, указывается индекс того стандарта, взамен которого введен настоящий его шифр по «Классификатору общесоюзных стандартов», наименование организации, разработавшей стандарт, даты утверждения и срока введения стандарта, а также его цена

К стандартам близки *межотраслевые и отраслевые технические документы* (технические условия, нормали и другие нормативные документы). В отличие от стандартов, они имеют более ограниченную сферу действия и создаются тогда, когда отсутствуют стандарты на данную продукцию или же когда в силу особых условий производства и эксплуатации возникает необходимость установления специальных требований к этой продукции. С 1960 г. установлены следующие виды технических условий: межреспубликанские (МРТУ), республиканские (РТУ), отдельных предприятий или их групп (ТУ). Нормали устанавливаются на узлы и детали машин, механизмов и приборов, а также на технологическую оснастку. Они определяют конструкцию, размеры, материалы и другие технические данные, необходимые для изготовления деталей и узлов машин. В настоящее время установлены следующие виды нормалей: межотраслевые (МН), отраслевые (ОН) и заводские (Н). К технической документации относятся и руководящие технические материалы (РТМ), выпускаемые отдельными учреждениями и организациями. В отличие от технических условий и нормалей их применение не является обязательным. В них излагаются рекомендуемые методы расчетов, сборки и сварки, способы применения защитных покрытий и т. п.

За рубежом, наряду со стандартами, широко издаются *спецификации*, представляющие собой коммерческие документы, содержащие описания технических требований к продуктам, материалам, процессам или видам обслуживания. Спецификации устанавливают требования, предъявляемые покупателями и потребителями к изготовителям. Зарубежные стандарты, как и советские, различаются по индексам, которые обычно состоят из сокращенного названия выпускающей стандарт организации, номера серии и года утверждения стандарта. Индекс «ASA. B 18,5 : 1952», например, означает: «ASA» — «American Standards Association», «B» — «Машиностроение», 18 — «Болты и гайки», 5 — серийный номер, 1952 — год утверждения стандарта. В капиталистических странах стандарты и примыкающая к ним нормативная литература издаются правительственными и частными организациями и, как правило, имеют лишь рекомендательный характер. В настоящее время стандарты играют важную роль в развитии науки, техники и торговли; они широко используются учеными, исследователями, инженерами, строителями и другими специалистами.

*Описания изобретений* к авторским свидетельствам и патентам (или иначе «патентная литература») являются одним из важных документальных источников научно-технической информации. Они позволяют проследить историю любого изобретения или открытия, а также выяснить современное состояние и направление развития технической мысли в соответствующей области. Современное ознакомление с патентной литературой предотвращает возможность повторного изобретения или разработки уже известных машин и методов. Необходимо указать, что сами авторские свидетельства и патенты являются лишь юридическими документами и не содержат никакой научно-технической информации. Они хранятся в одном экземпляре у изобретателя или патентообладателя и удостоверяют его права на сделанное изобретение.

В СССР основной формой охраны авторских прав изобретателя является авторское свидетельство. Оно обеспечивает бессрочное сохранение авторского права изобретателя на изобретение, а также право на получение вознаграждения и пользование всеми льготами, предоставляемыми в нашей стране изобретателям и рационализаторам. В соответствии с природой социалистического строя, при котором средствами производства владеет государство, авторское свидетельство предоставляет государству исключительное право на использование изобретения.

В капиталистических странах выдаются лишь патенты, предоставляющие их владельцам право в течение 15—17 лет использовать данное изобретение. В СССР патенты обычно выдаются изобретателям, являющимся гражданами иностранных государств.

Как уже было сказано, патентная литература — это не сами авторские свидетельства или патенты, а описания изобретений

к ним, составляемые в органах регистрации изобретений и размножаемые типографским способом. Советские описания изобретений (иначе «патентные описания») издаются в виде отдельных листков или брошюр и содержат следующие выходные сведения: классификационный индекс, обозначающий предмет изобретения по принятой в стране системе классификации патентов; номер авторского свидетельства или патента; указание вида издания; имя изобретателя или патентообладателя; название изобретения; дату заявки изобретения и номер заявочного свидетельства; указание учреждения, в котором заявлено изобретение; дату опубликования изобретения или ссылку на номер «Бюллетеня изобретений», в котором оно опубликовано.

В тексте описания дается краткое изложение существа изобретения, указываются области его применения, характеризуются известные средства решения данной задачи и перечисляются чертежи отдельных элементов изобретения. Затем следует полное описание изобретения, в котором приводятся характерные черты и сущность нового устройства, машины, изделия или химического вещества, перечисляются основные узлы и детали с объяснением их назначения и взаимодействия, а также ссылки на соответствующие позиции в чертежах. Патентное описание заканчивается так называемой «патентной формулой» — кратким изложением предмета изобретения с выделением того наиболее существенного, что считается изобретением и является объектом правовой защиты.

В ряде зарубежных стран патентное описание имеет некоторые специфические особенности. В патентных описаниях США, например, очень детально излагается предмет изобретения, тогда как в Великобритании это делается лишь в нескольких пунктах. В ФРГ при составлении патентной формулы, помимо изложения элементов новизны, указывается сущность изобретения, а во Франции, Бельгии и ряде других стран патентная формула вообще не составляется, так как экспертиза на новизну в этих странах не производится.

Патентная литература очень обширна: ежегодно в мире выпускается свыше 300 тыс. патентных описаний.

*Технические каталоги* (они называются также промышленными или торговыми каталогами) — это перечни изделий, выпускаемых промышленным предприятием (объединением) или продаваемых торговой организацией. Обычно они содержат технические описания изделий, которые сопровождаются иллюстрациями и чертежами. Иногда в технических каталогах указывается стоимость изделий. Технические каталоги могут содержать сведения о промышленной продукции одного вида и назначения или о продукции различной номенклатуры, выпускаемой предприятиями одного треста, министерства или одной фирмы. В них дается не только техническая характеристика изделий, но и обозначение типов или

моделей этих изделий, описание их конструкций и областей применения. По техническим каталогам можно изучать существующее оборудование с целью его приобретения, более рационального использования или усовершенствования.

Каталоги фирм капиталистических стран принято называть *фирменными каталогами*. Они издаются отдельными фирмами, концернами и торговыми объединениями в целях рекламы. В них не раскрываются конструктивные особенности рекламируемого оборудования и часто не даются его существенные технико-экономические показатели. Несмотря на это, иностранные фирменные каталоги служат ценным источником информации о новой технике.

Каталоги публикуются в виде листовок, брошюр, книг, альбомов, бюллетеней и серийных изданий. Особенно популярен тип каталога-справочника, содержащего сравнительные данные по многим видам однотипной продукции. Число их огромно; несмотря на отсутствие библиографического учета этих изданий, некоторые специалисты считают, что ежегодно публикуется до 0,5 млн. различных технических каталогов. Технические каталоги широко используются инженерами-проектировщиками, конструкторами и технологами, экономистами, изобретателями, научными работниками.

Поскольку в каталогах часто не указываются цены на оборудование и материалы, то приходится еще пользоваться отдельно издаваемыми прейскурантами и ценниками. *Прейскуранты* — это документы, имеющие временное значение, так как указываемые в них цены непрерывно меняются. Прейскуранты, издаваемые в СССР, содержат единые цены на материалы, оборудование и изделия, действительные в пределах всей страны или отдельных союзных республик. Кроме цен, в них иногда включаются основные технические характеристики оборудования, а также ссылки на стандарты или технические условия. Некоторые прейскуранты, помимо технических характеристик, содержат изображения и краткие описания машин или изделий и поэтому могут до некоторой степени заменять соответствующие технические каталоги. Как правило, они используются для составления смет, проектов и оформления расчетов между предприятиями.

С 1961 г. большинство прейскурантов издает Бюро цен Госплана СССР. К ним выпущено два перечня: «Перечень прейскурантов оптовых цен на промышленную продукцию, тарифов на перевозки и услуги, сдаточных, закупочных и оптовых цен на сельскохозяйственную продукцию» и «Перечень прейскурантов государственных розничных цен на продовольственные и промышленные товары». Прейскуранты, которые издаются частными фирмами в зарубежных странах, сильно отличаются от прейскурантов, издаваемых в нашей стране. Зарубежные прейскуранты

содержат только перечни цен и служат непосредственным дополнением к техническим каталогам.

К перечисленным видам изданий относят и материалы технической информации и обмена опытом, в которых освещаются новые достижения в области промышленного производства и техники. Они отличаются большим разнообразием в содержании и оформлении и издаются в виде отдельных книг и их серий, брошюр, периодических изданий и листовок. В качестве примера можно назвать *информационные листки*, которые обычно выпускаются сериями. Их тираж, как правило, не превышает нескольких сотен экземпляров. Информационные листки издаются территориальными ЦБТИ и центральными отраслевыми институтами информации и используются для обмена информацией между научно-исследовательскими институтами, конструкторскими бюро, проектными организациями и заводами с целью устранения параллелизма и дублирования в их работе. Эти листки содержат краткое описание (объемом до $^1/_4$ авторского листа) методов инженерных расчетов, новой технологии изготовления изделий, методов регулировки и настройки, проверки и испытаний. Информационные листки выпускаются по частным техническим решениям, имеющим самостоятельное значение еще до окончания всего исследования или конструкторской разработки. В них приводятся электрические и кинематические схемы отдельных элементов, узлов и блоков, их конструкции. а также характеристики и свойства разработанных материалов.

Большинство других материалов технической информации относятся либо к неопубликованным первичным документам (научно-технические отчеты, информационные карты), либо ко вторичным документам (информационные карточки). Эти виды документов рассматриваются в соответствующих разделах книги.

### *Неопубликованные документы*

Большое значение как источники научной информации имеют неопубликованные документы. Это — лабораторные журналы, научно-технические отчеты, информационные карты, а также отдельные данные, таблицы, чертежи, схемы, не вошедшие в опубликованные отчеты о научно-исследовательских и опытно-конструкторских работах. Некоторые материалы, обычно считающиеся произведениями печати, могут быть неизданными. Среди них можно назвать материалы научных конференций, диссертации, переводы и т. д.

*Научно-технические отчеты* (или отчеты о результатах законченных научно-исследовательских и опытно-конструкторских работ) служат важным источником научно-технической информации, необходимым для дальнейшего развития науки и техники, внедрения их достижений в производство и для усовершенствования

технологических процессов. Методика составления технических отчетов и их структура специфичны для каждой отрасли науки и техники, для отдельных стран и даже учреждений. Однако и в этих документах можно выделить некоторые общие элементы.

Обычно научно-технический отчет начинается с аннотации или реферата, в котором кратко излагаются задача исследования и полученные результаты. Затем следует введение, в котором кратко освещаются достижения отечественной и зарубежной науки и техники по данному вопросу. В тексте отчета излагается постановка задачи, приводится формулировка технического задания, анализируются известные методы и способы решения задачи, а также дается обоснование принятого решения, необходимые расчеты и результаты экспериментов. Отчет заканчивается выводами, в которых проводятся сопоставление и анализ теоретических и экспериментальных данных, полученных в процессе исследования. Затем следует заключение, где результаты работы оцениваются и сравниваются с требованиями технического задания, а также намечаются пути использования этих результатов. К научно-техническому отчету обычно прилагается обзор использованных источников. Иногда такой обзор служит первой частью самого отчета, в конце которого помещается тематический указатель использованной литературы.

Этот вид документов получил широкое распространение в послевоенные годы. В настоящее время в США только правительственными учреждениями ежегодно выпускается свыше 100 тыс. научно-технических отчетов [28].

Некоторые зарубежные специалисты предпочитают научно-технические отчеты журнальным статьям, поскольку в них полнее отражаются методика и техника проведения эксперимента. Научно-технические отчеты не считаются публикациями в полном смысле, хотя некоторые из них размножаются типографским способом; они имеют хождение среди ограниченного круга лиц и учреждений и, как правило, не подвергаются библиографическому учету. Сведения о технических отчетах почти не распространяются по традиционным библиографическим каналам. Поэтому службы научной информации должны обращать особое внимание на этот вид документов, в которых часто содержатся очень важные факты и сведения.

*Информационные карты* служат в СССР основной формой представления научно-технической, производственной и экономической информации предприятиями и организациями в центральные отраслевые органы научно-технической информации. По материалам информационных карт центральные информационные

[28] Science, government, and information. A report of the President's Science Advisory Committee. Washington, U. S. Government Printing Office, 1963, стр. 19.

органы издают библиографическую, реферативную и обзорную информацию о достижениях отечественных научно-исследовательских и проектно-конструкторских институтов и бюро, высших технических учебных заведений, университетов, предприятий, строек, совхозов, колхозов и других организаций. Карта составляется непосредственно на месте в одном экземпляре по стандартному образцу и подписывается руководителем учреждения или организации.

В информационных картах приводятся сведения о законченных научно-исследовательских и опытно-конструкторских работах, их этапах и частных технических решениях, имеющих самостоятельное значение; о новых изделиях и продуктах, внедренных в производство; о новых сортах сельскохозяйственных культур и породах скота; о новых технологических процессах, агрономических и зоотехнических приемах, а также о новых материалах, химикатах и т. д.

*Препринтами* (или предварительными публикациями) принято называть отдельные оттиски статьи или какой-либо другой работы, изготовленные типографским способом до ее официального опубликования в очередном номере периодического издания и предназначенные для рассылки ограниченному кругу заинтересованных лиц. Многие периодические издания и сборники, наряду с текстом основного издания, с того же набора печатают отдельные оттиски статей, передаваемые их авторам, которые рассылают эти оттиски своим коллегам. Такая практика существует уже давно. Она позволяет существенно ускорить доведение научных идей и фактов до сведения специалистов. Система отдельных оттисков подготовила возникновение препринтов, достоинство которых состоит в том, что они опережают официальную публикацию научной работы иногда на несколько месяцев.

Одной из разновидностей препринтов являются издания, которые выпускаются небольшим тиражом и предназначаются для обсуждения, проводимого с целью подготовки окончательного текста в соответствии с полученными критическими замечаниями. Эти издания носят самый разнообразный характер. Примерами могут служить тексты докладов и выступлений на конференциях и совещаниях, словники и макеты отдельных томов энциклопедий и справочников, подготовительные материалы к терминологическим словарям, спискам дескрипторов, тезаурусам, предварительные издания учебников и руководств, а также другие аналогичные материалы.

К неопубликованным документам следует отнести и переводы статей из иностранных журналов и сборников. Известно, что в США и Великобритании наиболее ценные советские журналы целиком переводятся на английский язык. Некоторые научно-технические журналы зарубежных стран переводятся на русский язык и в Советском Союзе. Некоторые переводы выпускаются

в виде тематических сборников и в сущности ничем не отличаются от книг и периодических изданий. Однако значительная часть переводов научно-технической литературы оформляется в виде машинописных рукописей. На обложке переводов на русский язык, выполняемых в нашей стране, обычно указывается название учреждения, в котором сделан перевод, порядковый номер перевода, дата и заглавие. На титульном листе, если он имеется, указывается автор в русской транскрипции и оригинальном написании, заглавие оригинала переведенной статьи и источник, из которого она взята. Из неопубликованных переводов создаются фонды — общесоюзный и отраслевые. Всесоюзный институт научной и технической информации, Государственная публичная научно-техническая библиотека, а также отраслевые институты информации и библиотеки координируют работу по переводам и издают каталоги и списки выполненных переводов.

Особое место занимают *диссертации и авторефераты* к ним. Диссертацией называется научное исследование, представленное на соискание ученой степени, автореферат — изложение основных положений диссертации, составленное автором для предварительного ознакомления с ними научной общественности. В СССР диссертации обычно не публикуются, но подвергаются строгой библиографической регистрации. Таким образом, хотя они и существуют в единичных экземплярах, но изложенные в них идеи и факты считаются официально введенными в научный оборот. Авторефераты диссертаций публикуются ограниченным тиражом (100—150 экземпляров), но обладают всеми правами произведения печати, хотя и помечаются иногда грифом «На правах рукописи». В СССР, США, Великобритании, Франции, ФРГ и других странах издаются текущие библиографии диссертаций.

К неопубликованным документам могут быть причислены также материалы, подготовленные к печати, т. е. *рукописи и корректуры научных работ*. Из-за длительности полиграфического процесса и по другим причинам удельный вес таких документов иногда достигает $^1/_3$ всей научно-технической литературы, опубликованной по любому вопросу в течение двух-трех последних лет [29].

По сравнению с произведениями печати неизданные материалы обладают рядом особенностей, которые приходится учитывать в научно-информационной деятельности. Эти документы существуют в единичных экземплярах. Они не всегда апробированы авторитетным научным коллективом (редколлегией, редакционным советом и т. п.) и могут не иметь всех необходимых выходных сведений (заглавия, подзаголовочных и надзаголовочных данных, указаний о месте и времени подготовки документа). Иногда такие

[29] Г. К о б л а н с. Новые методы и техника сообщения сведений «Бюлл. ЮНЕСКО для библиотек», 1957, т. 11, № 7, стр. 174.

документы не брошюруются и подготавливаются в виде отдельных листков. Рукописные и машинописные документы труднее читать и рèпродуцировать.

В заключение необходимо отметить, что выявление и сбор неопубликованных документов связаны с большими трудностями, так как сведения о таких документах не могут быть получены по традиционным библиотечно-библиографическим каналам, а их сбор, обработка, хранение и поиск требуют применения особой методики.

## ВТОРИЧНЫЕ ДОКУМЕНТЫ И ИЗДАНИЯ

Выше уже указывалось, что вторичные документы представляют собой конечные результаты аналитико-синтетической и логической переработки первичных научных документов. Если взять за основу степень такой переработки первичных документов, то вторичные документы можно условно расположить в следующий логический ряд: справочная литература, обзоры, рефераты и аннотации, материалы экспресс-информации, библиографические указатели, каталоги и картотеки.

### *Справочная литература*

Справочная литература предназначена для быстрого получения каких-либо сведений научного, прикладного или познавательного характера. В ней содержатся результаты теоретических обобщений, основные научные факты, математические и физико-химические величины, материалы производственного характера, сопровождающиеся большим количеством таблиц, графиков, чертежей и формул. Среди справочной литературы наибольшее значение имеют общие и отраслевые энциклопедии, производственные справочники, толковые словари, дву- и многоязычные словари, а также библиографические словари и справочники.

*Энциклопедиями* называются справочные издания, содержащие наиболее существенную (подробную или краткую) информацию по всем или по отдельным областям знания и практической деятельности. По содержанию они делятся на общие (универсальные) и отраслевые, по структуре — на алфавитные и систематические, по объему — на настольные (1—4 тома), малые (5—12 томов) и большие (в несколько десятков томов). Обычно энциклопедии составляются на основе тематического плана, по которому весь объем энциклопедии распределяется между различными научными дисциплинами или их разделами, и словника — списка всех слов, на которые подготавливаются отдельные статьи. Статьи в энциклопедии могут быть разного типа и объема. Наиболее обширными являются статьи-обзоры, охватывающие большие

темы. Однако большинство статей в энциклопедии составляют справки, содержащие определение и основную информацию о том или ином предмете: его теории, истории и т. д. В кратких статьях-толкованиях дается лишь определение и иногда происхождение слова. Наконец, в энциклопедиях имеется значительное число статей-отсылок, поясняющих, что данный материал следует искать под другим словом. Большое значение имеют списки литературы, помещенные в конце статей: они дают возможность читателю более углубленно изучать вопрос, по которому в энциклопедии дана лишь основная информация. Ко многим энциклопедиям составляются алфавитно-предметные указатели, которые повышают справочную ценность соответствующей энциклопедии в несколько раз.

В СССР издаются следующие общие энциклопедии. «Большая Советская Энциклопедия» (БСЭ) вышла в двух изданиях. Первое издание БСЭ, завершенное в 1947 г., состоит из 65 основных и одного дополнительного тематического тома «СССР». Второе издание БСЭ было выпущено в 1949—1958 гг. и содержит 49 основных и 2 дополнительных тома. Новые сведения публикуются регулярно в «Ежегодниках БСЭ». Энциклопедия снабжена двухтомным алфавитно-предметным указателем. «Малая Советская Энциклопедия» (МСЭ) опубликована в трех изданиях: первое издание в 10 томах выпускалось с 1928 по 1931 г., второе издание в 11 томах — с 1933 до 1947 г. и третье издание в 10 томах — с 1958 до 1961 г. К последнему изданию МСЭ выпущен отдельным томом алфавитно-предметный указатель. Общие настольные энциклопедии представлены «Энциклопедическим словарем» в трех томах (1953—1955 гг.), рассчитанным на широкие круги читателей. В 1963—1964 гг. вышло новое двухтомное издание этого словаря.

Отраслевые энциклопедии отличаются от универсальных не только тематикой, но и более ограниченным назначением: они, как правило, рассчитаны на специалистов и научных работников. Энциклопедии и энциклопедические словари издаются в СССР по многим отраслям знания. В качестве примеров можно назвать «Техническую энциклопедию», «Сельскохозяйственную энциклопедию», «Большую медицинскую энциклопедию», «Энциклопедию элементарной математики», «Физический энциклопедический словарь», «Геологический словарь», «Краткую химическую энциклопедию», «Географическую энциклопедию», «Философский словарь», словарь «Автоматизация производства и промышленная электроника», краткую энциклопедию «Атомная энергия», «Энциклопедический словарь-справочник по животноводству», «Советскую историческую энциклопедию», «Краткую литературную энциклопедию» и др.

Разумеется, невозможно перечислить все наиболее значительные энциклопедические издания, даже если ограничиться

только универсальными. Сведения о них приведены в специальных библиографических обзорах: по отечественным энциклопедиям хорошим путеводителем служит книга И. М. Кауфмана «Русские энциклопедии» (М., 1960), а по зарубежным — работы И. В. Гудовщиковой «Общие зарубежные энциклопедии» (Л., 1963) и (на английском языке) Р. Л. Коллисона «Энциклопедии, их история на протяжении веков. Библиографический справочник» (R. L. Collison. Encyclopaedias, their history throughout the ages. A bibliographical guide. New York, 1962).

Энциклопедии играют большую роль в научно-информационной деятельности. Однако пользоваться ими нужно критически. Прежде всего необходимо помнить, что подготовка энциклопедий в социалистических и капиталистических странах производится с разных идейных позиций. Желательно брать сведения из последних по времени изданий и проверять их по разным энциклопедиям, учитывая при этом характер и назначение каждого издания. Необходимо также уточнять, обозначает ли год издания, указанный на титульном листе, действительно год нового издания энциклопедии или просто год ее перепечатки.

*Производственные справочники* содержат сведения, которые могут потребоваться в работе специалистам определенного профиля.

К числу таких справочников относится, например, справочник-спутник «Хютте», перевод которого с немецкого языка переиздавался в СССР 15 раз. Издаются также специализированные производственные справочники, например, «Справочник металлиста», «Справочник химика», «Краткий справочник химика», «Краткий физико-технический справочник», «Справочник технолога-машиностроителя»; тематические справочники по материалам и оборудованию, по вопросам экономики производства, организации и охраны труда, по технике безопасности и другим вопросам, например, «Справочник по новой технике в сельском хозяйстве». Выпускаются также нормативные и квалификационные справочники.

*Толковые словари* дают краткие объяснения значений терминов той или иной отрасли науки и техники. В научно-информационной деятельности используются также политехнические словари, например, «Краткий политехнический словарь». Сведения о словарях подобного рода приведены в библиографическом указателе И. М. Кауфмана «Терминологические словари» (М., 1961).

К литературе рассматриваемого типа примыкают справочники констант, определители минералов, растений и животных, атласы чертежей и карт, альбомы оборудования, сборники формул, таблиц и графиков и т. п.

*Двуязычные и многоязычные словари* содержат слова и термины одного языка с их эквивалентами на другом языке. Двуязычные словари могут быть общими или специальными, т. е. содержа-

щими термины из определенной области. Многоязычные словари обычно бывают только специальными. В таких словарях каждому термину ставятся в однозначное соответствие его эквиваленты на других языках. Реже многоязычные словари составляются из нескольких двуязычных с общими указателями. Двуязычные и многоязычные словари оказывают большую помощь при чтении и переводе на русский язык научно-технической литературы. Сведения о специальных языковых словарях можно почерпнуть из «Библиографии многоязычных научных и технических словарей», изданной ЮНЕСКО (Bibliography of interlingual scientific and technical dictionaries. 4th ed. Paris, Unesco, 1960).

В биографических словарях приводятся данные о жизни, деятельности и трудах выдающихся ученых, инженеров, врачей, агрономов, общественных деятелей и т. д. Такие словари особенно полезны при изучении истории науки, техники и производства. Примером может служить «Биографический словарь деятелей естествознания и техники» в двух томах (М., 1958—1959). Отечественные словари этого типа перечислены в библиографическом указателе И. М. Кауфмана «Русские биографические и библиографические словари» (М., 1955).

### *Обзоры*

Ежегодные обзоры развития исследований в различных областях науки и техники известны сравнительно давно. Лондонское физическое общество, например, публикует «Report on progress in physics» с 1934 г.

За последнее десятилетие обзоры заняли прочное место в системе вторичных научных документов. Они обычно выпускаются в виде сборников, издаваемых под редакцией крупнейших специалистов той или иной области. Обзоры позволяют ученым и инженерам следить за основным направлением развития соответствующей научной дисциплины.

Всесоюзный институт научной и технической информации издает такие обзоры в серии «Итоги науки и техники». Обзоры выпускаются и другими научными учреждениями и организациями, а также отдельными издательствами, публикующими научную литературу.

В более широком плане обзор представляет собой вторичный научный документ, в котором обобщаются сведения, содержащиеся в соответствующих первичных документах. Обзоры обычно охватывают материалы по одной теме и за определенный период времени. Различают два вида обзоров — аналитический и реферативный.

*Аналитический обзор* является результатом всестороннего анализа первичных научных документов, содержит аргументи-

рованную оценку материала и дает обоснованные рекомендации относительно перспектив развития и использования соответствующих достижений науки и техники. Такого рода обзоры могут рассматриваться как научно-исследовательские работы и нередко составляют важную часть этих работ.

*Реферативный обзор* характеризуется меньшей глубиной анализа и строится на основе извлечения и систематизации определенных данных из первичных научных документов. Он обычно содержит сведения о новейших достижениях отечественной и зарубежной науки и техники и позволяет судить о состоянии и тенденциях развития определенной отрасли или отдельного вопроса. Реферативные обзоры распространены значительно шире, чем аналитические обзоры, и по сравнению с тематическими библиографическими указателями или сборниками рефератов обладают рядом преимуществ. Ознакомление с правильно составленным обзором не требует последующего обращения к первичным документам; обзор отличается значительно большей целенаправленностью; он знакомит специалиста с достижениями в смежных областях знания и производства, помогает при выборе направления и методов разработок в определенной области, а также при уточнении профиля комплектования.

Аналитические и реферативные обзоры по определенному вопросу могут составляться регулярно и даже периодически, но чаще всего — это отдельные работы, охватывающие материалы за определенный промежуток времени. Обзоры могут размножаться типографским способом, средствами малой полиграфии или же остаются неопубликованными материалами, предназначенными для внутреннего использования.

### *Реферативные журналы*

*Реферативный журнал* — это периодическое издание, в котором помещаются рефераты и аннотации первичных научных документов по какой-либо отрасли знания или производства. В настоящее время реферативный журнал является важнейшим видом вторичных информационных изданий.

В «Реферативном журнале ВИНИТИ» помещаются вторичные научные документы трех видов: рефераты, аннотации и библиографические справки. *Рефераты* составляют на большинство работ, освещаемых в реферативном журнале. Рефераты должны отличаться лаконичностью стиля, в них широко применяются условные сокращения часто повторяющихся терминов. В реферате излагаются цели, методы, основные теоретические предпосылки и результаты работы, приводятся цифровые данные, формулы, таблицы, графики и схемы. В конце реферата указывается число иллюстраций, таблиц и библиографических ссылок в реферируемой работе, место ее выполнения и составитель реферата. *Ан-*

*нотации* в предельно краткой форме сообщают о содержании и характере работы и обычно составляются на второстепенные публикации. *Библиографические справки* помещаются непосредственно перед каждым рефератом и аннотацией, но могут рассматриваться и как самостоятельные вторичные документы. Они содержат порядковый номер публикации в реферативном журнале, индекс Универсальной десятичной классификации (УДК) и полное библиографическое описание первичного документа. Для работ на иностранных языках в начале описания приводится перевод заглавия на русский язык, а в конце — указание на язык оригинала. Каждая серия или отдельный выпуск реферативного журнала снабжается алфавитными указателями — предметным и авторским; в некоторых случаях помещаются формульные и иные специальные указатели.

В настоящее время реферативный журнал выполняет две основные функции: а) является средством распространения информации о новых достижениях науки и техники и б) служит информационно-поисковой системой, позволяющей производить поиск документов по определенным вопросам. Нужно признать, что пока он выполняет обе эти функции не вполне удовлетворительно: первую — из-за недостаточной оперативности (реферат появляется в реферативном журнале не менее чем через четыре месяца после опубликования соответствующей статьи), вторую — из-за несовершенства справочного аппарата и отсутствия кумуляции (при поиске литературы за продолжительный отрезок времени приходится просматривать большое число годовых указателей). Для существенного повышения эффективности реферативного журнала необходимо значительное сокращение сроков его подготовки, дальнейшее расширение охвата мировой научно-технической литературы, повышение научного уровня рефератов, совершенствование системы указателей к реферативным журналам, а также резкое сокращение сроков их составления путем применения соответствующих средств механизации.

Количество реферативных изданий исчисляется сотнями и продолжает расти. В СССР основные реферативные журналы издает ВИНИТИ. Их характеристика дана в главе «Организация научно-информационной деятельности».

### *Экспресс-информация*

Выше было указано, что одним из основных недостатков современных реферативных журналов является слишком большой разрыв во времени между опубликованием первичного научного документа и его отражением в реферативном журнале. Стремясь к тому, чтобы ученые и инженеры как можно быстрее узнавали о выходе новых работ по их специальности, информационные центры, библиографические учреждения и специальные библио-

теки прибегают к параллельному выпуску более оперативных информационных изданий.

*Экспресс-информация* — довольно широкий и неопределенный термин, обозначающий различные виды информационных изданий. Это может быть комплект печатных карточек, обычный библиографический бюллетень, упрощенный библиографический указатель, в котором опущены некоторые элементы описания (например, заглавие статьи дано только на языке читателя и не дано на языке оригинала), сборники оглавлений только что опубликованных журналов и т. п. Их объединяет лишь та особенность, что все они созданы с целью ускорения информации о новых публикациях.

В последнее время название «экспресс-информация» стало применяться для обозначения бюллетеней, включающих переводы (сокращенные или полные) публикаций о важнейших изобретениях, открытиях или исследованиях, информацию о которых нужно срочно довести до сведения работников науки и техники. Именно такой характер носит «Экспресс-информация» ВИНИТИ.

Появление такого вида вторичных изданий, как экспресс-информация, обусловлено главным образом недостатками современных реферативных журналов. Существование экспресс-информации, наряду с реферативными журналами, значительно усложняет информационную систему, делает ее более громоздкой и дорогостоящей. По-видимому, экспресс-информация в разных ее формах — явление временное. Необходимость в подобных изданиях исчезнет, когда будут найдены более совершенные средства вторичной информации.

### *Библиографические указатели*

*Библиографические указатели* — это составленные по определенному плану печатные или рукописные перечни литературы, которые облегчают ориентировку в этой литературе, пропагандируют ее и содействуют ее лучшему использованию. Библиографические указатели являются одним из важнейших видов вторичных документов и изданий и играют большую роль в научной информации.

Область знания, предметом которой является учет, описание, систематизация и качественный анализ произведений печати или рукописей, а также составление библиографических указателей, называется библиографией. Термин «библиография» многозначен и используется также для обозначения самих библиографических указателей и пособий, а также их совокупности. Поэтому при рассмотрении различных видов библиографических указателей можно говорить о тех или иных видах библиографии.

По целевому назначению библиография делится на учетно-регистрационную, информационную и рекомендательную. *Учетно-регистрационная библиография* обеспечивает исчерпывающую регистрацию литературы по отраслям знаний, предметам или видам изданий. *Информационная библиография* знакомит широкие круги читателей с новейшей литературой. *Рекомендательная библиография* призвана выполнять задачу активного руководства чтением и помощи определенным категориям читателей в правильном выборе литературы.

По содержанию и виду отражаемых произведений библиография подразделяется на общую и специальную. *Общая библиография* учитывает произведения независимо от их содержания и вида, а *специальная библиография* — только произведения по определенной отрасли знания или только произведения определенного вида. Кроме того, библиографию принято подразделять на *универсальную* (т. е. учитывающую литературу независимо от ее содержания) и *отраслевую* (т. е. учитывающую литературу по определенной отрасли знания или предмету).

По времени издания отражаемых произведений библиография может быть текущей, ретроспективной и проспективной. *Текущая библиография* учитывает произведения по мере их выпуска в свет, *ретроспективная* — за определенный промежуток времени, предшествующий моменту составления библиографии, а *проспективная* — отражает произведения, подготавливаемые к изданию, но еще не вышедшие в свет.

По месту издания отражаемой литературы различают международную, государственную и краевую библиографию. *Международная библиография* учитывает литературу, независимо от того, в какой стране она издана. Такие библиографии часто составляются международными организациями с привлечением библиографов разных стран. *Государственная библиография* отражает произведения печати, вышедшие на территории одной страны. В СССР этим термином принято обозначать учетно-регистрационные библиографические издания, выпускаемые специальными государственными учреждениями (например, «Летописи», издаваемые книжными палатами). За рубежом в этом значении иногда употребляют термин «национальная библиография». Кроме того, национальная библиография часто понимается как библиография произведений, написанных авторами одной народности, или же на одном языке. *Краевая библиография* учитывает местную печать, т. е. издания, вышедшие на территории края, области, района или какой-либо другой местности.

По широте географического района, о котором говорится в учитываемой литературе, библиография подразделяется на страноведческую и краеведческую. *Страноведческая библиография* охватывает литературу об одной или нескольких странах, а крае-

ведческая — литературу о крае, области, районе или какой-либо местности.

По виду отражаемой литературы различают: библиографию книг, библиографию периодических изданий и статей, библиографию официальных изданий, библиографию трудов научных учреждений и организаций, библиографию трудов конференций и съездов, библиографию диссертаций, библиографию библиографий (так называемую библиографию второй степени). Необходимо отметить, что существуют библиографии трех степеней, которые принято называть соответственно библиографиями, или библиографическими указателями (первая степень), библиографиями библиографий, или путеводителями по библиографиям (вторая степень), и библиографиями третьей степени.

По читательскому назначению библиографию делят на *научную* (т. е. рассчитанную на ученых и квалифицированных специалистов) и *массовую* (т. е. предназначенную для широкого круга читателей).

По источнику составления различают *первичную* библиографию, составленную непосредственно на основе просмотра самих произведений, и *вторичную*, составленную на основе других библиографических указателей, каталогов и картотек.

Существует также много других видов библиографий: по полноте охвата литературы — исчерпывающая и выборочная; по дробности отражения литературы — учитывающая произведения или издания в целом (не имеет специального названия) или учитывающая лишь части произведений или изданий (*аналитическая*); по методам раскрытия содержания литературы — *описательная*, *формальная*, *аннотированная*, *реферативная*; по принципу расположения записей — алфавитная, систематическая и предметная; по внешней форме самой библиографии — *книжная*, *периодическая*, *внутрикнижная*, *прикнижная*, *пристатейная*, *микрофильмированная* и т. д.

Приведенное перечисление видов библиографии не является исчерпывающим. Оно должно дать лишь общее представление о многообразии библиографических источников и помочь работникам органов научной информации в анализе конкретных библиографических указателей. Количество библиографических указателей и пособий огромно. Только в нашей стране ежегодно выпускаются тысячи названий различных библиографий и более 150 специальных периодических изданий библиографического характера. Если же учитывать библиографии, выходящие в зарубежных странах, то приходится говорить о десятках и сотнях тысяч названий. Для правильной ориентировки в библиографии необходимо хотя бы в общих чертах познакомиться с основными методами библиографической работы, знать наиболее важные библиографические пособия, которые могут быть использованы для поиска литературы по интересующей теме.

В данной монографии невозможно рассмотреть даже основные библиографические издания по всем отраслям знания. Поэтому здесь придется ограничиться лишь упоминанием наиболее важных текущих и ретроспективных библиографий второй степени и общих библиографий.

*Библиография библиографий* (или библиография второй степени) является ключом ко всем библиографическим источникам. Основным руководством в этой области может служить выпущенное Государственной публичной библиотекой им. М. Е. Салтыкова-Щедрина методическое и справочное пособие И. К. Кирпичевой «Библиография в помощь научной работе» (Л., 1958). В этом капитальном труде освещается методика библиографической работы и характеризуются основные отечественные и зарубежные библиографические источники сведений о литературе по всем отраслям знания (универсальная библиография) и по отдельным отраслям знания (отраслевая библиография).

Источником текущей информации о советских библиографических изданиях является ежегодник «Библиография советской библиографии», выпускаемый Всесоюзной книжной палатой. В этом ежегоднике отражаются все виды библиографических указателей, а также работы по теории и методике библиографии. Зарубежные библиографические издания описаны в книгах Г. Г. Кричевского «Общие библиографии зарубежных стран» (М., 1962), К. Р. Симона «История иностранной библиографии» (М., 1963), К. Уинчелл «Путеводитель по справочникам» с дополнениями за 1950—1952, 1953—1955, 1956—1958 гг. (содержит сведения более чем о 8 тыс. общих и отраслевых библиографий и справочников) [30], Л. Мальклес «Библиографические источники» [31], а также в книге В. Тотока, Р. Вейтцеля и К. Веймана «Руководство по библиографическим справочникам» [32]. Во «Всемирной библиографии библиографий» Т. Бестермена [33] учтено около 85 тыс. отдельных библиографических изданий, выпущенных с конца XV в. до 1954 г.

*Текущая государственная библиография в СССР* издается Всесоюзной книжной палатой в виде «Летописей». Источником сведений о книгах служит еженедельный библиографический указатель «Книжная летопись», в котором регистрируются не только книги и брошюры, но и стандарты, типовые проекты, технические

---

[30] C. M. Winchell. Guide to reference books. 7th ed. Chicago, 1951.

[31] L. N. Malcles. Sources du travail bibliographique, v. 1—3. Geneve—Lille, 1950—1958.

[32] W. Totok, R. Weitzel, K.-H. Weimann. Handbuch der bibliographischen Nachschlagewerke. 3. erw., völlig neu bearb. Aufl. Frankfurt a/M., 1966.

[33] T. Besterman. World bibliography of bibliographies and of bibliographical catalogues, calendars, abstracts, digests, indexes and the like, v. 1—4. 3rd ed. Geneve, 1955—1956.

Библиотечный каталог — это указатель имеющихся в библиотеке произведений, составленный для ориентировки читателей и библиотекарей в библиотечном фонде. Каталоги советских библиотек служат основным средством раскрытия содержания библиотечного фонда, пропаганды книг и руководства чтением.

Каталоги имеют много общего с библиографическими указателями, поскольку и те, и другие являются перечнями произведений печати. Отличительная особенность библиотечного каталога состоит в том, что он отражает состав фонда определенной библиотеки, тогда как в библиографический указатель включаются произведения, независимо от места их нахождения. Каталоги различаются по способу группировки библиографических описаний, по видам отражаемых произведений печати, по широте охвата библиотечных фондов, по языку произведений печати, по назначению, по форме выполнения самих каталогов, а также по другим признакам.

По способу группировки описаний каталоги подразделяются на алфавитные, систематические и предметные. В *алфавитном каталоге* описания произведений печати, независимо от их содержания, располагаются в алфавитном порядке фамилий авторов и названий учреждений, а также заглавий произведений, принадлежащих более чем трем авторам или изданных без указания автора. Подобное расположение материала позволяет получать ответы на два главных вопроса: 1) имеется ли в библиотеке определенное произведение, автор и заглавие которого известны; 2) какие произведения данного автора, вышедшие в виде отдельных изданий, имеются в библиотеке. Если в алфавитный каталог включается достаточное число добавочных описаний на лиц, участвовавших в подготовке тех или иных изданий, то в таком каталоге можно найти ответ на некоторые второстепенные вопросы (например, в создании каких произведений печати определенное лицо принимало участие в качестве составителя, редактора, иллюстратора, переводчика и т. п.).

Алфавитный каталог является одним из основных справочно-библиографических средств информационного органа или библиотеки, так как дает ответ на вопрос о наличии в фонде определенного издания. Большинство читателей, приходя в библиотеку, уже знают, какие книги им нужны. По алфавитному каталогу они выясняют, имеются ли интересующие их книги в библиотеке, и узнают шифры, по которым эти книги расставлены в книгохранилище. Вся внутренняя работа библиотеки — комплектование, обработка и хранение фондов, обслуживание читателей, справочно-библиографическая и массовая работа — возможна лишь на базе алфавитного каталога.

Алфавитный каталог служит также единственным средством раскрытия авторского состава библиотечного фонда. Он способствует пропаганде произведений основоположников марксизма-

ленинизма, классиков отечественной и зарубежной науки, выдающихся деятелей современности во всех областях общественной жизни. Поэтому было бы неправильным считать алфавитный каталог формальным и отводить ему второстепенную роль среди каталогов других видов.

Для облегчения подбора литературы по содержанию и удовлетворения тематических запросов читателей информационные органы и библиотеки ведут систематические и предметные каталоги. В *систематическом каталоге* описания произведений печати группируются по отраслям знания, с которыми связано их содержание. Внутри отраслевого раздела эти описания сначала группируются по классам, внутри каждого класса — по подклассам и т. д. Структура систематического каталога библиотеки находится в прямой зависимости от применяемой в этой библиотеке системы библиотечно-библиографической классификации. Различные системы таких классификаций рассматриваются в главе «Аналитико-синтетическая переработка документов». Систематические каталоги раскрывают тематическое содержание справочно-информационных фондов и поэтому служат важным средством научной информации. Кроме того, систематические каталоги играют важную роль в работе библиотек. Группировка произведений печати по отраслям знания значительно облегчает руководство чтением и оказание помощи в выборе книг. Во внутренней работе библиотеки систематические каталоги используются при комплектовании, в справочно-библиографической деятельности и в других библиотечных процессах.

Из сказанного следует, что при помощи систематического каталога всегда можно установить, какие документы по данному вопросу имеются в справочно-информационном фонде. Однако если вопрос достаточно узок, а читатель не является специалистом в соответствующей области, то ему не так легко определить, к какому подразделению классификации, использованной для составления каталога, относится интересующий его вопрос. Для облегчения работы с систематическим каталогом к нему составляется *алфавитно-предметный указатель* (или «ключ»). В алфавитно-предметном указателе приводятся в общем алфавитном порядке все рубрики систематического каталога, по которым в фонде имеется литература. Кроме того, в указатель могут входить рубрики, отсутствующие в соответствующей схеме классификации, а также широко употребляемые синонимы терминов, содержащихся в этой схеме.

В *предметном каталоге* произведения также раскрываются по их содержанию, но описания произведений группируются по признаку предмета, о котором в них идет речь. Слово или словосочетание, обозначающее какой-либо один предмет (понятие) и используемое для выражения предметного содержания документов или их частей, в библиотековедении называется *предметной*

*рубрикой* (или предметным заголовком). В предметном каталоге все предметные рубрики располагаются в общем алфавитном порядке. Предметные рубрики могут быть простыми и сложными.

Сложной предметной рубрикой считается такая, которая обозначает группу простых, называемых в этом случае *подрубриками* (например, Машинный перевод — автоматический синтез, алгоритмы, библиографии по МП, история, конференция, лексическая омонимия и полисемия, программирование, семантический анализ, словари, формализация записи, эксперименты, язык-посредник).

Эффективность предметного каталога в значительной степени зависит от правильной формулировки предметных рубрик. Поэтому важно, чтобы все словосочетания, используемые в качестве предметных рубрик, начинались со слов, которые являются наиболее существенными в данных конкретных условиях. Для этого при формулировке рубрик часто прибегают к инверсии, т. е. к соответствующей перестановке слов в рубрике. Так, например, понятиям «Исторические архивы» и «Личные архивы» в предметном каталоге будут соответствовать рубрики «Архивы, исторические» и «Архивы, личные».

Предметные каталоги особенно удобны для отражения литературы по прикладным наукам — техническим, сельскохозяйственным и медицинским. Поэтому они распространены главным образом в специальных библиотеках. Научные универсальные библиотеки обычно ведут предметные каталоги только на часть своих фондов, а массовые библиотеки вообще не составляют предметных каталогов. Особое распространение получили предметные каталоги в советских медицинских библиотеках, которые отражают в них и отечественную и зарубежную литературу и включают в эти каталоги, наряду с описаниями книг, также и описания журнальных статей.

Перечисленные виды каталогов являются основными по способу группировки описаний. Для отдельных видов произведений печати (например, стандартов, ведомственных технических документов, описаний изобретений к авторским свидетельствам и патентам и т. п.) применяются *номерные каталоги* (их часто называют также нумерационными каталогами). Произведения печати указанных видов чаще разыскиваются по их индексам или номерам, а не по заглавиям или авторам. Поэтому в каталоге целесообразно располагать описания этих произведений печати не в алфавитном порядке, а в порядке их номеров. Таким образом, номерной каталог выполняет для этих видов произведений печати роль алфавитного каталога. Иногда для пропаганды актуальной литературы по какому-либо вопросу или проблеме создаются *тематические картотеки*. По структуре они напоминают систематический каталог, но не могут его заменить, так как выбор вопросов или проблем, отражаемых в тематических карто-

теках, носит временный характер, а отражение литературы в таких картотеках всегда бывает выборочным. На практике иногда встречаются комбинации из основных видов каталогов. Так, в некоторых специальных библиотеках ведутся *систематическо-предметные каталоги*, в которых основные разделы расположены по схеме классификации (т. е. в систематическом порядке), а входящие в эти разделы предметные рубрики — в алфавитном порядке. В США широко распространен *словарный каталог*, представляющий собой комбинацию алфавитного и предметного каталогов. Опыт показывает, что механическое объединение нескольких принципов в одном каталоге снижает его эффективность. Поэтому в советских библиотеках стремятся к созданию системы каталогов, которая позволяет наиболее полно с минимальными затратами средств раскрыть библиотечный фонд.

Остальные виды каталогов различаются по второстепенным признакам. По виду отражаемых произведений печати различаются *каталоги карт*, *каталоги периодических изданий*, *каталоги специальных видов технической литературы*; по широте охвата фондов — *сводные каталоги* (каталоги фондов нескольких или многих библиотек), *генеральные каталоги* (включающие все фонды библиотеки), *центральные каталоги* (отражающие фонды всех филиалов и отделов в основном вузовских библиотек), *каталоги отдельных фондов*, *отделов и филиалов* библиотеки; по языку произведений печати — *каталоги литературы на русском языке*, *каталоги литературы на языках других народов СССР*, *каталоги литературы на иностранных языках*.

По назначению каталоги бывают *читательскими* и *служебными*, а по форме выполнения — карточными и книжными. *Карточный каталог* — основной вид каталогов, применяемых в советских библиотеках. Карточки для таких каталогов имеют международный формат 125×75 *мм*, изготовляются из плотной бумаги и хранятся по 800—1000 экз. в ящиках специальных шкафов. Карточки в каталожном ящике закрепляются при помощи съемного стержня, проходящего через специальные отверстия, которые пробиты у нижнего края карточек. Главное достоинство карточного каталога — возможность включения новых описаний без нарушения принятого порядка расстановки описаний в каталоге. Это позволяет поддерживать карточные каталоги на уровне современного состояния справочно-информационного или библиотечного фонда. Вместе с тем карточный каталог обладает и рядом недостатков. Он громоздок, быстро изнашивается, для его просмотра требуется много времени. Каталоги в форме книги (*книжные каталоги*) внешне напоминают библиографические указатели. Они компактны, удобны в работе, меньше изнашиваются. Средства современной полиграфии позволяют размножать книжные каталоги в большом числе экземпляров, что несомненно повышает степень их использования. Однако такие

каталоги можно дополнять новыми библиографическими описаниями через сравнительно большие промежутки времени, при их переиздании. Указанными достоинствами и недостатками этих двух форм каталогов можно объяснить наблюдаемую в крупных зарубежных библиотеках тенденцию к замене части карточных каталогов печатными каталогами в книжной форме и к сохранению карточных каталогов лишь для новых поступлений.

Для быстрого изготовления печатных каталогов в настоящее время используются специальные устройства, автоматически переснимающие на рулонную фотопленку текст с верхнего края 80-колонных перфокарт. При помощи наборно-пишущей машины типа «Verityper» вдоль верхнего края такой перфокарты печатаются одна или две строки библиографического описания, рубрики указателя или какой-либо другой текст. Остальное поле перфокарты используется для пробивки отверстий, соответствующих закодированным характеристикам текста. Массив перфокарт, составляющий каталог, сортируется в том или ином порядке (алфавитном, систематическом или предметном). Упорядоченный массив затем пропускается через специальные устройства «Fotolist», «Composo-line» или «Listomatic», автоматически фотографирующие текст с заданной кратностью уменьшения. Устройства «Fotolist Model 270» и «Composo-line» работают со скоростью 7200 перфокарт/час, а камера «Listomatic» — со скоростью 13 800 перфокарт/час. Полученная в этих камерах негативная фотопленка используется затем для изготовления офсетных печатных форм.

Из других форм выполнения каталогов и картотек следует отметить *плоские обозримые картотеки* (например, типа «Кардекс»), которые широко применяются в зарубежных информационных органах и библиотеках для регистрации периодических изданий. Плоская обозримая картотека представляет собой металлический шкаф с выдвижными плитами, на которых шарнирно закреплены (в лежачем положении) карточки. Нижний край каждой карточки, на котором записаны основные сведения об издании, остается открытым. Плоские обозримые картотеки в известной мере сочетают достоинства каталогов, выполненных в карточной и книжной формах.

Значение каталогов и картотек как источников научной информации приближается к значению библиографических указателей. Печатные каталоги таких национальных книгохранилищ, как Библиотека Британского музея и Библиотека конгресса США, по существу являются полными перечнями англо-американской книги. В СССР начато издание сводного каталога книг четырех крупнейших книгохранилищ — Государственной библиотеки СССР им. В. И. Ленина, Всесоюзной книжной палаты, Государственной публичной библиотеки им. М. Е. Салтыкова-Щедрина и Библиотеки Академии наук СССР. Этот сводный каталог будет

наиболее полной библиографией русской книги с начала книгопечатания до наших дней. Каталоги многих специализированных библиотек по своему значению не уступают соответствующим отраслевым библиографиям. В двух отношениях они даже превосходят библиографические указатели: они непрерывно пополняются описаниями все новых поступлений и дают непосредственный адрес самого издания на библиотечной полке, тогда как литературу, которая указана в библиографии, нужно еще разыскивать в библиотеках.

Библиографические картотеки как источник информации приобретают особое значение в связи с тем, что ряд крупных библиотек и библиографических учреждений выпускает готовые печатные карточки с описаниями и аннотациями на новые книги и статьи. В СССР карточки на отечественные книги и журнальные статьи выпускает Всесоюзная книжная палата, на зарубежные книги и статьи по естественным наукам — Всесоюзная государственная библиотека иностранной литературы, по технике — Государственная публичная научно-техническая библиотека. Печатные картотеки по отдельным разделам науки и техники издает и распространяет ВИНИТИ. Карточки обычно снабжены классификационными индексами, которые позволяют расставлять описания в систематическом порядке и использовать картотеку для тематического подбора литературы. Наряду с библиографическими карточками, выпускаются также информационные карточки, которые издаются центральными реферативными органами на основе технических отчетов, информационных листков и других справочно-информационных материалов. Они содержат индекс УДК, библиографическое описание и аннотацию и предназначены для картотек справочно-информационных фондов.

### *Вспомогательные указатели*

*Вспомогательные указатели* являются важной составной частью многих первичных и вторичных документов и изданий. Речь идет об указателях авторов, заглавий, названий учреждений, персоналий, о предметных и иных указателях к монографиям, сборникам, энциклопедиям, библиографиям и другим документам и изданиям. Вспомогательные указатели (их иногда называют «ключами») позволяют отыскивать в научных документах и изданиях сведения, которые не выявлены группировкой материала в основном тексте. Эти указатели значительно повышают научную и справочную ценность научных документов и изданий.

В качестве примера можно указать, что вспомогательные указатели к библиографическим изданиям позволяют отразить учтенную в них литературу по нескольким разным признакам. Если материал в библиографии расположен в систематическом порядке, то авторский и предметный указатели дают возможность отыски-

вать по этой библиографии необходимые документы не только по отраслям знаний, но и по отдельным вопросам, предметам, понятиям, именам.

Принципы построения вспомогательных указателей в основе своей ничем не отличаются от описанных выше принципов построения каталогов и библиографических указателей. По способу группировки материала они также подразделяются на алфавитные, систематические и предметные. Однако методика составления вспомогательных указателей имеет свою специфику, зависящую не только от вида указателя, но и от вида документа или издания, к которым они составлены. Наиболее распространенными из вспомогательных указателей являются именные и предметные.

*Именной указатель* обычно содержит перечень фамилий с инициалами или полными именами, расположенный в алфавитном порядке. После фамилий и имен указываются номера страниц основного текста или библиографических описаний в библиографическом издании. Если данное лицо является автором какого-либо произведения, то иногда в указателе рядом с его фамилией приводится заглавие произведения. Во вспомогательных указателях к библиографиям нередко приводятся фамилии и названия, отобранные только по одному какому-либо признаку. Тогда мы имеем дело с *указателем авторов*, *указателем коллективов* (названий учреждений и организаций, от имени которых издается тот или иной документ), *указателем персоналий* (фамилий и имен лиц, о которых сообщается в документах). Имена, принадлежащие гражданам иностранных государств, по большей части приводятся в оригинальном написании и группируются отдельно от имен, написанных по-русски, образуя самостоятельный алфавитный ряд. Иногда иностранные имена приводятся в именном указателе в русской транскрипции и объединяются в общем алфавитном ряду с русскими именами. Этот метод в последнее время все шире практикуется в советских изданиях, хотя он обладает существенным недостатком. Поскольку нормы транскрипции иностранных имен в русском тексте разработаны не очень четко, одни и те же имена часто оказываются в разных местах указателя по-разному написанными.

*Предметный указатель* объединяет под каждой предметной рубрикой ссылки на номера страниц или библиографических описаний, в которых нашел отражение определенный вопрос. Предметные указатели очень важны в научных документах и изданиях, так как предметная группировка материала в основном тексте этих документов и изданий применяется редко. Но даже и в изданиях с предметным расположением материала, например в энциклопедических словарях, вспомогательный предметный указатель помогает глубже раскрывать содержание документов.

Основной чертой вспомогательного предметного указателя является то, что он отражает лишь наиболее существенные во-

просы, предметы, понятия каждого документа. В этом он ничем не отличается от предметных каталогов, алфавитно-предметных указателей к систематическим каталогам, библиографических указателей с предметным расположением материала. Часто встречаются предметные указатели, где рубрики отражают материал в особом аспекте, наиболее характерном для содержания документа или издания, к которым они составлены, например в географическом, геологическом и биологическом. Эти вспомогательные указатели обычно получают специальные названия типа «Указатель географических названий», «Указатель геологических терминов», «Указатель по видам животных» и т. п., хотя по существу они являются предметными.

*Алфавитно-предметный указатель* является комбинацией именного и предметного вспомогательных указателей. В нем все рубрики — фамилии и имена лиц, заглавия произведений, названия учреждений и предметов — расположены в общем алфавитном порядке. Чаще всего такие указатели составляются к систематическим каталогам. Значительно реже алфавитно-предметные указатели составляются к научным документам и изданиям. Это практикуется лишь в тех случаях, когда какие-либо сведения, например фамилии и имена лиц или названия учреждений, настолько немногочисленны, что выделять их в самостоятельный указатель нецелесообразно.

В *тематическом указателе* в отличие от всех рассмотренных видов вспомогательных указателей, для которых характерно алфавитное расположение материала, рубрики сгруппированы в логическом порядке. Он встречается редко, главным образом в изданиях рекомендательного характера, рассчитанных на широкий круг потребителей информации. При помощи тематического указателя удобнее всего раскрывать содержание научных изданий по крупным проблемам или многоаспектным вопросам.

### *Новые виды указателей*

Одним из основных недостатков традиционных указателей, которые были рассмотрены в предыдущих разделах, является сравнительно длительный срок их подготовки. Это обусловлено большой трудоемкостью составления таких указателей, а также тем, что их составление плохо поддается механизации и автоматизации.

Кроме того, в научно-информационной практике возникают такие требования, удовлетворить которые невозможно при помощи традиционных указателей (часто, например, появляется необходимость узнать, в работах каких авторов была использована та или иная публикация).

Все это приводит к появлению новых видов указателей, отличительной особенностью которых служит то, что они составляются

Если при составлении пермутационного указателя справа или слева от входного ключевого слова в строке остается незанятое место, то туда из другой части строки может быть перенесен (переставлен) не умещающийся в ней микроконтекст. Переносить можно не только целые слова, но и любые части слова (даже одну букву). Когда переносимый микроконтекст целиком не умещается в свободной части строки, то он обрывается на последнем полностью уместившемся слове. Усечение микроконтекста отмечается знаком «+». Следовательно, при составлении указателей такого типа производится перестановка слов в заглавиях. Именно поэтому эти указатели называются пермутационными (от англ. to permute — переставлять, менять порядок).

При составлении пермутационного указателя возможны случаи, когда переносы слов заглавия из одной части строки в другую невозможны из-за отсутствия в ней незанятых мест. В таких случаях заглавие публикации соответствующим образом усекается. Разрыв слова допускается лишь тогда, когда в строке остается не менее четырех букв.

Пермутационный указатель имеет следующие основные недостатки:

1) заглавие не всегда в должной мере раскрывает содержание публикации;

2) в качестве самостоятельных предметных рубрик в указателе используются синонимичные ключевые слова. Поэтому при информационном поиске приходится припоминать все синонимичные ключевые слова и просматривать в указателе их микроконтексты. Это не только существенно замедляет процесс поиска, но и может быть причиной значительных потерь информации;

3) в указателе не выявляются и не отмечаются родо-видовые связи, существующие между ключевыми словами заглавий. Это также приводит к существенным потерям информации при поиске;

4) переход от выделенного ключевого слова к библиографическому описанию публикации, в заглавие которой входит это ключевое слово, производится при помощи достаточно сложного кода, который нужно каждый раз запоминать или записывать. Это не может не затруднять процесс поиска;

5) указатель имеет сравнительно большой объем.

Однако перечисленные недостатки пермутационных указателей с избытком компенсируются тем, что указатели такого типа составляются в основном при помощи электронных цифровых вычислительных машин (ЭЦВМ). Это позволяет выпускать пермутационные указатели в очень короткие сроки — всего за несколько дней. Именно этим можно объяснить быстро растущую популярность пермутационных указателей среди непосредственных потребителей научной информации — ученых и инженеров. В настоящее время в мире выпускаются сотни указателей такого типа.

Первые образцы пермутационных указателей были составлены при помощи ЭЦВМ в 1958 г. Их разработали одновременно и, по-видимому, независимо друг от друга специалисты американских фирм System Development Corp. и International Business Machines Corp.[34]

Однако из литературы известно, что при помощи алфавитных счетно-перфорационных машин указатель такого типа составлялся в США еще в 1953 г.[35]

Один из основных создателей пермутационного указателя — американский инженер Г. П. Лун, работавший в фирме International Business Machines Corp. Именно он внес наибольший вклад в дело популяризации таких указателей, их широкого внедрения в научно-информационную практику.

В 1960 г. редакция реферативного журнала «Chemical Abstracts» (США) приступила к подготовке пермутационного указателя «Chemical Titles», который начал распространяться по подписке с января 1961 г. За четыре года существования этот указатель получил всеобщее признание во всем мире.«Chemical Titles» — двухнедельный библиографический бюллетень сигнального типа, в котором раскрывается содержание статей примерно из 600 наиболее важных журналов по химии и химической технологии. Необходимо отметить, что в этих журналах помещается не менее 50% всех публикаций, освещаемых на страницах реферативного журнала «Chemical Abstracts» (всего в этом журнале реферируется около 9000 периодических и продолжающихся изданий).

Бюллетень «Chemical Titles» состоит из трех частей:

1) собственно пермутационный указатель;
2) библиографический указатель индексируемых статей;
3) авторский указатель.

В каждом номере бюллетеня раскрывается содержание в среднем 2850 статей.

Подготовка бюллетеня «Chemical Titles» производится следующим образом. Сначала библиографические описания статей, подлежащих включению в бюллетень «Chemical Titles», пробиваются на перфокартах. Все описания даются на английском языке. Заглавия статей и имена их авторов пробиваются на отдельных перфокартах. На каждой перфокарте пробивается также адресный код публикации и ее порядковый номер в соответствующем выпуске журнала или продолжающегося издания. Когда наступает

---

[34] J. Citron, L. Hart, H. Ohman. A permutation index to the Preprints of the International Conference on Scientific Information. Santa Monica, System Development Corp., 1958; H. P. Luhn. Key-Word-In-Context Index for technical literature (KWIC Index). Yorktown Heights, IBM Advanced Systems Development Division, 1959.

[35] M. P. Veilleux. Permuted title word indexing procedures for a man-machine system. В кн.: «Machine indexing: progress and problems». Washington, The American University, 1962, стр. 77—111.

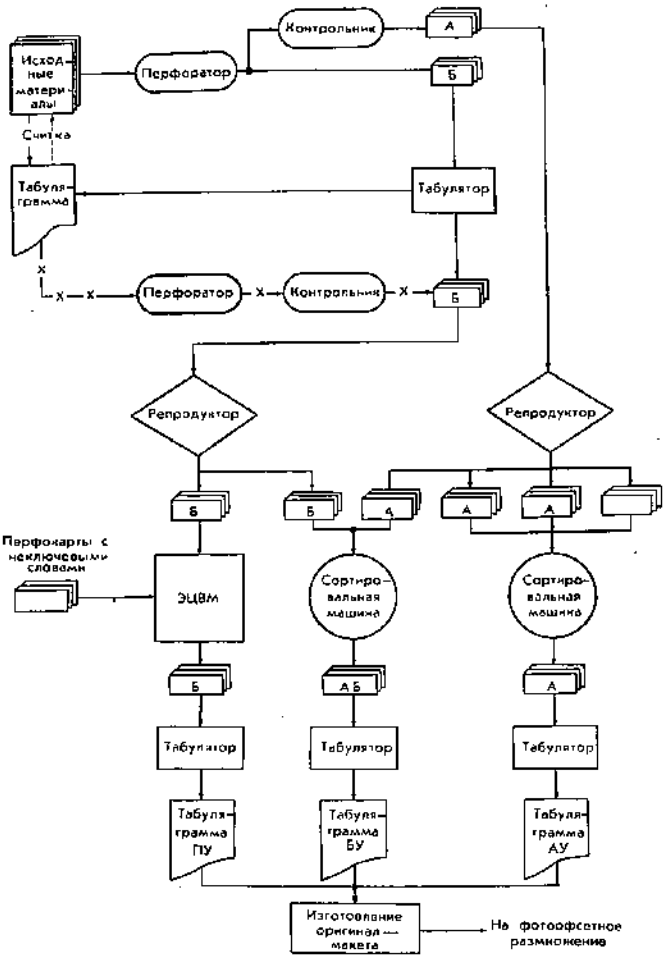


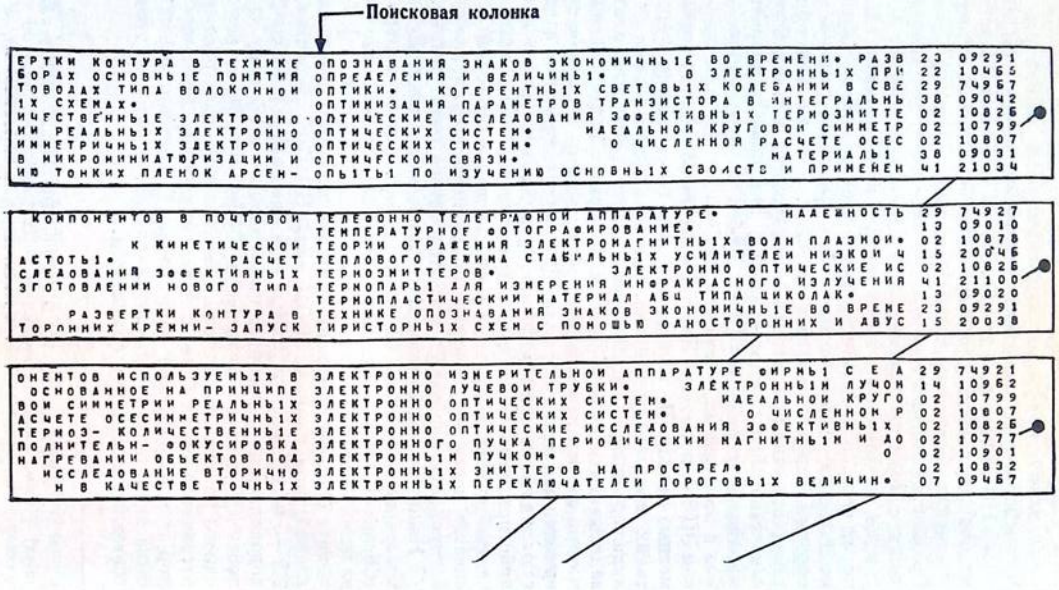

**Рис. 15. Примерная технологическая схема подготовки бюллетеня «Chemical Titles»**

срок выпуска очередного номера бюллетеня, накопленные массивы таких перфокарт передаются на машинную обработку. Примерная технологическая схема подготовки бюллетеня «Chemical Titles» показана на рис. 15. Из этой схемы видно, что ЭЦВМ используется лишь для составления собственно пермутационного указателя. Остальные части бюллетеня «Chemical Titles», а именно библиографический и авторский указатели, составляются при помощи алфавитных счетно-перфорационных машин. В ЭЦВМ подаются перфокарты с заглавиями статей, а также список неключевых слов. В настоящее время в этом списке насчитывается 1295 неключевых слов — предлогов, артиклей, союзов, некоторых наречий, вспомогательных глаголов, а также слов общего характера (типа «анализ», «отчет», «эксперимент» и т. п.). С этим списком сопоставляется каждое заглавие, введенное в оперативное запоминающее устройство ЭЦВМ. Любое слово заглавия, не включенное в этот список, считается ключевым. По этим ключевым словам ЭЦВМ, произведя необходимые перестановки, формирует предметные записи. На каждое заглавие составляется в среднем шесть таких предметных записей. Длина каждой записи — 60 знаков, включая интервалы. Конечные результаты всех этих операций записываются на перфокартах. В пермутационном указателе предметные записи должны быть расположены в алфавитном порядке выделенных ключевых слов. Для этого предметные записи подвергаются алфавитной сортировке, которая выполняется либо в ЭЦВМ, либо при помощи алфавитной сортировальной машины. Упорядоченный массив перфокарт с предметными записями далее пропускается через табулятор, и полученная табуляграмма после макетирования передается на фотоофсетное размножение.

Некоторое представление о количестве времени, затрачиваемого непосредственно на формирование одного номера бюллетеня «Chemical Titles», могут дать следующие цифры, относящиеся к начальному периоду выпуска этого бюллетеня (1961 г.). На создание номера бюллетеня «Chemical Titles» объемом 125 страниц (2850 заглавий) при помощи ЭЦВМ типа IBM-704 затрачивалось около 2,5—3 часов времени работы этой машины; еще около 2 часов расходовалось на перезапись информации с перфокарт на магнитную ленту и на печать табуляграмм. Таким образом, на все составление одного номера бюллетеня с готового массива перфокарт в 1961 г. затрачивалось не более 5 часов машинного времени [36]. Но с тех пор технологическая подготовка бюллетеня «Chemical Titles» была значительно усовершенствована, а расходы времени на его составление заметно сокращены.

---

[36] R. Freeman, M. Dyson. Development and production of «Chemical Titles», a current awareness index publication prepared with the aid of a computers. «Journal of Chemical Documentation», 1963, v. 3, № 1, стр. 16—20.




ЕРТКИ КОНТУРА В ТЕХНИКЕ ОПОЗНАВАНИЯ ЗНАКОВ ЭКОНОМИЧНЫЕ ВО ВРЕМЕНИ• РАЗВ 23 09291
БОРАХ ОСНОВНЫЕ ПОНЯТИЯ ОПРЕДЕЛЕНИЯ И ВЕЛИЧИНЫ• О ЭЛЕКТРОННЫХ ПРИ 22 10465
ТОВОДАХ ТИПА ВОЛОКОННОЙ ОПТИКИ• КОГЕРЕНТНЫХ СВЕТОВЫХ КОЛЕБАНИЙ В СВЕ 29 74967
ИХ СХЕМАХ• ОПТИМИЗАЦИЯ ПАРАМЕТРОВ ТРАНЗИСТОРА В ИНТЕГРАЛЬНЫ 38 09042
ИЧЕСТВЕННЫЕ ЭЛЕКТРОННО-ОПТИЧЕСКИЕ ИССЛЕДОВАНИЯ ЭФФЕКТИВНЫХ ТЕРМОЭМИТТЕ 02 10826
НИ РЕАЛЬНЫХ ЭЛЕКТРОННО ОПТИЧЕСКИХ СИСТЕМ• ИДЕАЛЬНОЙ КРУГОВОЙ СИММЕТР 02 10799
ИММЕТРИЧНЫХ ЭЛЕКТРОННО ОПТИЧЕСКИХ СИСТЕМ• О ЧИСЛЕННОЙ РАСЧЕТЕ ОСЕС 02 10807
В МИКРОМИНИАТЮРИЗАЦИИ И ОПТИЧЕСКОЙ СВЯЗИ• МАТЕРИАЛЫ 38 09031
ИЮ ТОНКИХ ПЛЕНОК АРСЕНИ- ОПЫТЫ ПО ИЗУЧЕНИЮ ОСНОВНЫХ СВОЙСТВ И ПРИМЕНЕН 41 21034

КОМПОНЕНТОВ В ПОЧТОВОЙ ТЕЛЕФОННО ТЕЛЕГРАФНОЙ АППАРАТУРЕ• НАДЕЖНОСТЬ 29 74927
ТЕМПЕРАТУРНОЕ ФОТОГРАФИРОВАНИЕ• 13 09010
К КИНЕТИЧЕСКОЙ ТЕОРИИ ОТРАЖЕНИЯ ЭЛЕКТРОМАГНИТНЫХ ВОЛН ПЛАЗМОЙ• 02 10878
АСТОТЫ• РАСЧЕТ ТЕПЛОВОГО РЕЖИМА СТАБИЛЬНЫХ УСИЛИТЕЛЕЙ НИЗКОЙ Ч 15 20046
СЛЕДОВАНИЙ ЭФФЕКТИВНЫХ ТЕРМОЭМИТТЕРОВ• ЭЛЕКТРОННО ОПТИЧЕСКИЕ ИС 02 10826
ЗГОТОВЛЕНИИ НОВОГО ТИПА ТЕРМОПАРЫ ДЛЯ ИЗМЕРЕНИЯ ИНФРАКРАСНОГО ИЗЛУЧЕНИЯ 41 21100
ТЕРМОПЛАСТИЧЕСКИЙ МАТЕРИАЛ АБЦ ТИПА ЦИКОЛАК• 13 09020
РАЗВЕРТКИ КОНТУРА В ТЕХНИКЕ ОПОЗНАВАНИЯ ЗНАКОВ ЭКОНОМИЧНЫЕ ВО ВРЕМЕ 23 09291
ТОРОННИХ КРЕМНИЙ- ЗАПУСК ТИРИСТОРНЫХ СХЕМ С ПОМОЩЬЮ ОДНОСТОРОННИХ И ДВУС 15 20038

ОНЕНТОВ ИСПОЛЬЗУЕМЫХ В ЭЛЕКТРОННО ИЗМЕРИТЕЛЬНОЙ АППАРАТУРЕ ФИРМЫ С Е А 29 74921
ОСНОВАННОЕ НА ПРИНЦИПЕ ЭЛЕКТРОННО ЛУЧЕВОЙ ТРУБКИ• ЭЛЕКТРОННЫМ ЛУЧОМ 14 10962
ВОЙ СИММЕТРИИ РЕАЛЬНЫХ ЭЛЕКТРОННО ОПТИЧЕСКИХ СИСТЕМ• ИДЕАЛЬНОЙ КРУГО 02 10799
АСЧЕТЕ ОСЕСИММЕТРИЧНЫХ ЭЛЕКТРОННО ОПТИЧЕСКИХ СИСТЕМ• О ЧИСЛЕННОМ Р 02 10807
ТЕРМОЭ- КОЛИЧЕСТВЕННЫЕ ЭЛЕКТРОННО ОПТИЧЕСКИЕ ИССЛЕДОВАНИЯ ЭФФЕКТИВНЫХ 02 10826
ПОЛНИТЕЛЬН- ФОКУСИРОВКА ЭЛЕКТРОННОГО ПУЧКА ПЕРИОДИЧЕСКИМ МАГНИТНЫМ И АО 02 10777
НАГРЕВАНИИ ОБЪЕКТОВ ПОД ЭЛЕКТРОННЫМ ПУЧКОМ• О 02 10901
ИССЛЕДОВАНИЕ ВТОРИЧНО ЭЛЕКТРОННЫХ ЭМИТТЕРОВ НА ПРОСТРЕЛ• 02 10832
Н В КАЧЕСТВЕ ТОЧНЫХ ЭЛЕКТРОННЫХ ПЕРЕКЛЮЧАТЕЛЕЙ ПОРОГОВЫХ ВЕЛИЧИН• 07 09467

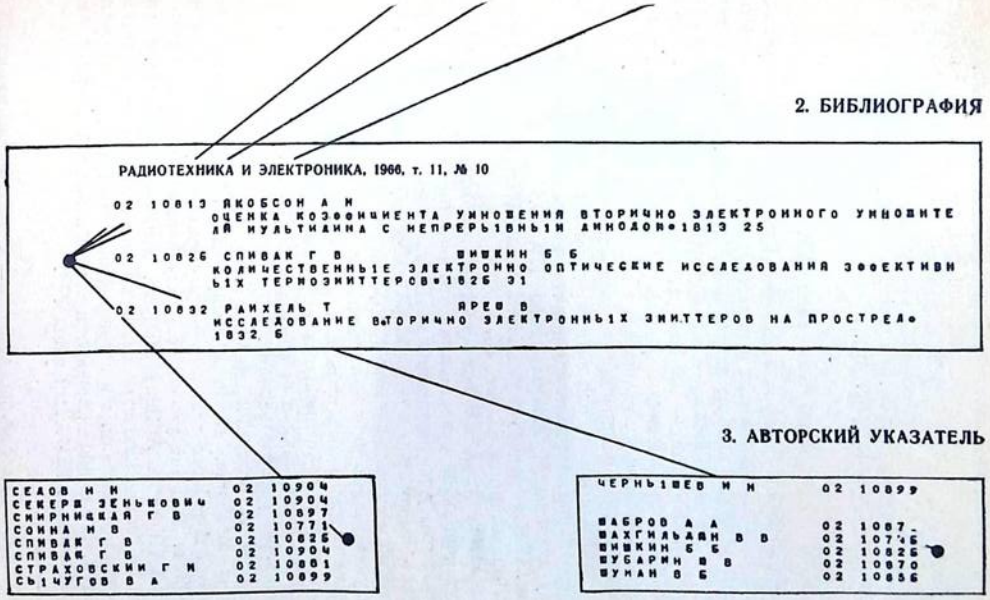



Рис. 16. Система указателей в бюллетене СИ ВИНИТИ «Электроника и ее применение»

Пермутационные указатели находят все более широкое применение и в советской научно-информационной практике. Они используются как справочный аппарат отдельных выпусков РЖ ВИНИТИ, например РЖ «Научная и техническая информация» (с 1967 г.), а также входят в состав некоторых выпусков сигнальной информации (СИ), например СИ ВИНИТИ по автоматике и радиоэлектронике (рис. 16).

### Указатели типа KWOC[37]

Библиографические указатели типа KWOC, как и пермутационные указатели, основаны на использовании в качестве предметных рубрик ключевых слов из заглавий публикаций. Однако в таких указателях каждое ключевое слово дается не в контексте, а за его пределами. Контекст — полное заглавие публикации — приводится под этим ключевым словом в виде предметной записи. Каждому выделенному ключевому слову приписан адресный шифр публикации (как в приведенном выше примере пермутационного указателя) или номер, под которым данная публикация отражена в реферативном журнале или другом каком-либо издании, к которому составлен указатель типа KWOC.

Рис. 17 иллюстрирует указатель типа KWOC к шести публикациям, приведенным на рис. 14.

Указатели типа KWOC обладают тем преимуществом перед пермутационными указателями, что они имеют привычную для читателя схему расположения материала и дают более обширный контекст, в котором употреблено выделенное ключевое слово. Недостатком указателей типа KWOC по сравнению с пермутационными указателями является значительно больший объем при равном числе отражаемых публикаций. В остальном указатели типа KWOC ничем не отличаются от пермутационных указателей; они также составляются при помощи ЭЦВМ и поэтому получают все большее распространение.

В 1963 г. редакция американского реферативного журнала «Applied Mechanics Reviews» выпустила алфавитно-предметный указатель подобного типа к годовому комплекту своего журнала за 1962 г. В этом указателе, получившем условное название WADEX (Word & Authors inDEX), в общем алфавитном ряду даются не только ключевые слова заглавий рефератов, аннотаций и рецензий, помещенных в журнале, но и имена авторов этих публикаций. Их имена включаются также в контекст, приводимый под выделенным ключевым словом [38].

---

[37] KWOC — Key-Word-Out-of-Context — Ключевые слова вне контекста.

[38] E. A. Ripperger a. o. WADEX (Word & Authors inDEX). A new tool in literature retrieving. «Applied Mechanics Reviews», 1963, v. 16, № 12, стр. 951—956.

| | |
|---|---|
| АННОТИРОВАНИЯ | СОБИ-5748-044 |
| О МЕТОДИКЕ АННОТИРОВАНИЯ ПРОИЗВЕДЕНИЙ ОСНОВОПОЛОЖНИКОВ МАРКСИЗМА-ЛЕНИНИЗМА | |
| АРХЕОЛОГИЧЕСКОЙ | НТИН-6308-013 |
| ИНФОРМАЦИОННЫЙ ПОИСК В ОБЛАСТИ АРХЕОЛОГИЧЕСКОЙ ЛИТЕРАТУРЫ | |
| БИБЛИОТЕК | БЮБИ-6206-285 |
| ОБОРУДОВАНИЕ НЕБОЛЬШИХ БИБЛИОТЕК | |
| ЕСТЕСТВЕННЫХ | НТИН-6308-021 |
| ОБ ОДНОМ ПОДХОДЕ К ТИПОЛОГИИ СЕМАНТИКИ ЕСТЕСТВЕННЫХ ЯЗЫКОВ | |
| ЗНАЧЕНИЙ | ВИМК-6103-022 |
| О КУЛЬТУРНОМ ЗНАЧЕНИИ МАШИННОГО ПЕРЕВОДА | |
| ИНФОРМАЦИОННЫЙ | НТИН-6308-013 |
| ИНФОРМАЦИОННЫЙ ПОИСК В ОБЛАСТИ АРХЕОЛОГИЧЕСКОЙ ЛИТЕРАТУРЫ | |
| КУЛЬТУРНОМ | ВИМК-6103-022 |
| О КУЛЬТУРНОМ ЗНАЧЕНИИ МАШИННОГО ПЕРЕВОДА | |
| ЛЕНИНИЗМА | СОБИ-5748-044 |
| О МЕТОДИКЕ АННОТИРОВАНИЯ ПРОИЗВЕДЕНИЙ ОСНОВОПОЛОЖНИКОВ МАРКСИЗМА ЛЕНИНИЗМА | |
| ЛИТЕРАТУРЫ | НТИН-6308-013 |
| ИНФОРМАЦИОННЫЙ ПОИСК В ОБЛАСТИ АРХЕОЛОГИЧЕСКОЙ ЛИТЕРАТУРЫ | |
| МАРКСИЗМА | СОБИ-5748-044 |
| О МЕТОДИКЕ АННОТИРОВАНИЯ ПРОИЗВЕДЕНИЙ ОСНОВОПОЛОЖНИКОВ МАРКСИЗМА ЛЕНИНИЗМА | |
| МАШИННОГО | ВИМК-6103-022 |
| О КУЛЬТУРНОМ ЗНАЧЕНИИ МАШИННОГО ПЕРЕВОДА | |
| МАШИНОМ | НТИН-6311-031 |
| ПОЗИЦИОННЫЙ ЭТАП СИНТЕЗА ПРИ МАШИННОМ ПЕРЕВОДЕ | |
| ОБОРУДОВАНИЕ | БЮБИ-6206-285 |
| ОБОРУДОВАНИЕ НЕБОЛЬШИХ БИБЛИОТЕК | |
| ПЕРЕВОДЕ | НТИН-6311-031 |
| ПОЗИЦИОННЫЙ ЭТАП СИНТЕЗА ПРИ МАШИННОМ ПЕРЕВОДЕ | |
| ПОЗИЦИОННЫЙ | НТИН-6311-031 |
| ПОЗИЦИОННЫЙ ЭТАП СИНТЕЗА ПРИ МАШИННОМ ПЕРЕВОДЕ | |
| ПОИСК | НТИН-6308-013 |
| ИНФОРМАЦИОННЫЙ ПОИСК В ОБЛАСТИ АРХЕОЛОГИЧЕСКОЙ ЛИТЕРАТУРЫ | |
| СЕМАНТИКИ | НТИН-6308-021 |
| ОБ ОДНОМ ПОДХОДЕ К ТИПОЛОГИИ СЕМАНТИКИ ЕСТЕСТВЕННЫХ ЯЗЫКОВ | |
| СИНТЕЗА | НТИН-6311-031 |
| ПОЗИЦИОННЫЙ ЭТАП СИНТЕЗА ПРИ МАШИННОМ ПЕРЕВОДЕ | |
| ТИПОЛОГИИ | НТИН-6308-021 |
| ОБ ОДНОМ ПОДХОДЕ К ТИПОЛОГИИ СЕМАНТИКИ ЕСТЕСТВЕННЫХ ЯЗЫКОВ | |
| ЯЗЫКОВ | НТИН-6308-021 |
| ОБ ОДНОМ ПОДХОДЕ К ТИПОЛОГИИ СЕМАНТИКИ ЕСТЕСТВЕННЫХ ЯЗЫКОВ | |

Рис. 17. Образец указателя ключевых слов из заглавий вне контекста (KWOC)

Объем указателя WADEX к реферативному журналу «Applied Mechanics Reviews» за 1962 г., содержащему в годовом комплекте около 8 тыс. публикаций, составил около 600 страниц. По ориентировочным данным, при наличии отлаженных программ для ЭЦВМ и массива перфокарт с занесенными на них именами авторов, заглавиями и номерами публикаций, на составление оригинал-макета указателя типа KWOC такого объема при помощи ЭЦВМ типа IBM-1401 требуется не более 20 часов машинного времени.

## Указатель «Index Medicus»

Большой интерес представляет также библиографический указатель «Index Medicus», выпускаемый Национальной медицинской библиотекой США (рис. 18). Библиотека была основана в 1836 г. и с 1879 г. выпускает этот предметный указатель отечественных и

**INDEX MEDICUS** ARTERIOSCLEROSIS

with benzine] Di Biasi S, et al
Folia Med (Napoli) 46:1094-103, Dec 63 (It)
[Ecgraphic findings in acute experimental poisoning with acrolein] Scorsone A, et al
Folia Med (Napoli) 46:1070-7, Dec 63 (It)
[Contribution to the study of filarial cardiopathy] Pannier R, et al Mal Cardiov 5:113-6, 1964 (It)
[Further observations on coronary sinus rhythm. Statistical study] Tomasi AM, et al
Minerva Cardioangiol 12:260-3, Jul 64 (It)
[Primary myocarditis in the adult caused by Coxsackie virus (group B, type 5) with plurirecurrent syncopal manifestations due to paroxysmal ventricular hyperactivity. Preliminary anatomo-clinical contribution] Pansini R, et al
Minerva Med 55:1641-55, 23 May 64 (It)
[Electrocardiographical study of 550 cases of typhoid infection treated with chloramphenicol and prednisone derivatives] Cristofani M
Osped Maggiore 59:163-76, Feb 64 (It)
[On the treatment of some abnormalities of cardiac rhythm in chronic fatigue in athletes] Venerando A, Florence] (Boncinelli E), Cossio M, et al
Riv Sper Freniat 88:Suppl:275-345, 31 Mar 64 (It)

### ARTERIES (A7)

Role of the autonomic nervous system in the nutrition of the products of conception. Effect of pelvic parasympathectomy on uterine and subplacental decidual blood flow. Toth A, et al
Fertil Steril 15:263-71, May-Jun 64
Stereological studies on several ducts and vessels by injection method of acrylic resin. 13. Arterial distribution of the upper lip in some mammals. Hanai H, Okajima. Folia Anat Jap 40:81-128, Apr 64
[On the absence of an artery in the umbilical cord] Gomori Z, et al
Gynaecologia (Basel) 157:177-90, 1964 (Ger)
[Aspects of metabolism of the arterial wall in animals treated with 5-hydroxytryptamine] Vernoni S
Arch Med Intern 15:158-63, Sep-Dec 63 (It)
[Hormones and fibrinolysis. Experimental investigations on the plasma and arterial walls]

Рис. 18. Фрагмент указателя «Index Medicus»

зарубежных статей по медицине. Он выходит ежемесячно с последующей ежегодной кумуляцией. Непрерывно возрастающий поток медицинской литературы сделал процессы подготовки и издания этого указателя сложными и трудоемкими. Уже в 1961 г. в каждом его выпуске регистрировалось более 10 тыс. статей. Ежегодное количество учитываемых статей возрастет со 140 тыс. в 1961 г. до 250 тыс. в 1969 г. За этот же период число названий расписываемых журналов предполагается увеличить с 14 тыс. до 22,5 тыс.

Для того чтобы справиться с этой задачей и улучшить качество указателя, библиотека создала автоматическую систему

MEDLARS (Medical Literature Analysis and Retrieval System — Система анализа и поиска медицинской литературы). В этой системе применена ЭЦВМ типа H-800 и специализированная фотонаборная машина типа GRACE, скорость работы которой достигает 300 знаков/сек. Использование автоматической системы позволило сократить срок подготовки ежемесячных выпусков с 22 до 5 дней и увеличить глубину индексирования документов в среднем с 2 до 10 предметных заголовков.

В указателе «Index Medicus» в качестве предметных записей используются полные библиографические описания объемом около 250 знаков каждая. Они распределяются по 6380 предметным рубрикам с коэффициентом дублирования одной и той же записи в разных рубриках (глубиной индексирования), достигающим 8—10. Каждый ежемесячный номер указателя достигает объема в 600 страниц и содержит до 14 тыс. библиографических описаний. Срок его подготовки, включая редактирование и тиражное размножение, не превышает 5 дней.

Все операции, выполняемые системой, разделяются на три группы. Подсистема ввода и преобразования включает получение журналов, книг и других документов, их описание и индексирование, прием читательских запросов, кодирование библиографических описаний и читательских запросов на перфоленту для ввода в машину. Эти операции выполняются квалифицированными библиографами. Подсистема хранения и поиска реализуется на ЭЦВМ, в которую информация с перфоленты после корректуры перезаписывается на магнитную ленту внешнего накопителя. Все сведения хранятся в записи на магнитную ленту и регулярно кумулируются для последующей выдачи по читательским запросам. Подсистема вывода при помощи фотонаборной машины преобразует сигналы магнитной ленты в буквы и цифры, фотографируемые на микрофильм, с которого печатается очередной номер «Index Medicus», а также библиографические указатели и списки по запросам читателей.

Система MEDLARS обеспечивает автоматическое составление текущих и ретроспективных списков подобного рода по индивидуальным и коллективным запросам. Она является хорошим примером комплексного подхода к решению проблем быстрого распространения библиографической информации и документального информационного поиска. Некоторые американские специалисты предлагают использовать этот опыт при создании национальной информационной системы США [39].

---

[39] Medical literature analysis and retrieval system. «Data Processing», 1965, v. 7, № 1, стр. 58—60; The MEDLARS story at the National Library of Medicine. Washington, 1965; W. K n o x. The changing role of libraries. «ALA Bulletin», 1965, v. 59, № 8, стр. 720—725.

## Указатели табличного типа

Указатели табличного типа (так называемые «scan-column indexes») представляют собой простейшие документальные информационно-поисковые системы, выполненные в виде книги [40]. В них применяется координатное индексирование, при котором основное смысловое содержание документов выражается в виде соответствующих перечней дескрипторов — слов и словосочетаний естественного языка, выбираемых из специально составленного для этой цели нормативного словаря (тезауруса). Перечень дескрипторов, характеризующих основное смысловое содержание документов, называется поисковым образом этого документа.

Процедура информационного поиска состоит в сопоставлении поисковых образов документов с данным поисковым предписанием — перечнем дескрипторов, характеризующих смысловое содержание информационного запроса. Документ считается отвечающим на информационный запрос, если его поисковый образ полностью или частично совпадает с поисковым предписанием. Для ускорения процедуры сопоставления поисковых образов документов с поисковыми предписаниями применяются различные технические средства, начиная от унитерм-карт и перфокарт ручного обращения и кончая быстродействующими ЭЦВМ. В указателях табличного типа эта процедура выполняется визуально.

На рис. 19 показан фрагмент указателя табличного типа. В первой колонке этого указателя записываются адресные шифры документов, а в последующих колонках против каждого адресного шифра приводятся коды дескрипторов, образующих поисковый образ этого документа. Код дескриптора состоит из номера колонки и символа (строчной или прописной буквы, какого-либо специального знака и т. п.), записываемого в этой колонке.

Информационный поиск в указателе табличного типа производится следующим образом. Сначала указатель просматривается по коду первого дескриптора, входящего в поисковое предписание. Если в определенном ряду колонки, номер которой входит в этот код, он содержится, то соответствующий ряд проверяется на вхождение в него других кодов поискового предписания. В случае совпадения всех кодов поискового предписания с кодами поискового образа данного ряда адресный шифр данного ряда указывает на местонахождение искомого документа.

Указатель такого типа уже более 35 лет успешно применяется в Патентном управлении Нидерландов. Он используется там при экспертизе на новизну патентных заявок по определенным разделам техники. В указателе табличного типа при информационном поиске по одному дескриптору может быть найдено до 1000 доку-

---

[40] J. O' Connor. The scan-column index. «American Documentation», 1962, v. 13, № 2, стр. 204—209.

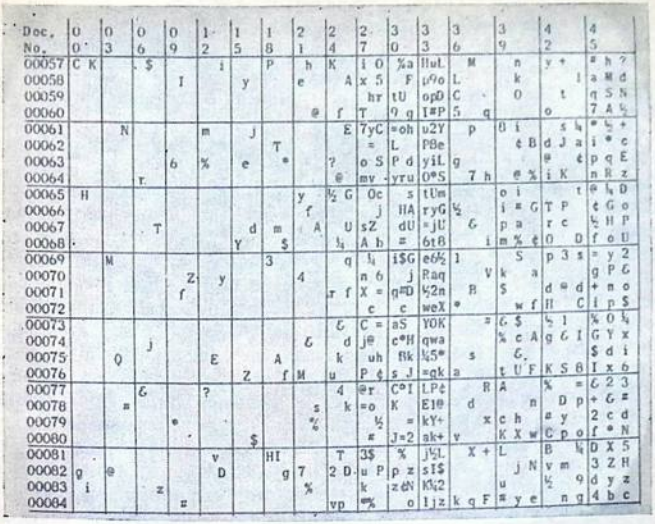


Рис. 19. Фрагмент табличного указателя

ментов в час, причем полнота поиска составляет около 75%.

В современных указателях табличного типа скорость просмотра в зависимости от плотности заполнения колонок указателя может колебаться в пределах от нескольких сотен до тысячи и более документов в минуту.

Основное достоинство указателей табличного типа состоит в том, что они могут быть получены как побочный выход информационно-поисковой системы, реализованной при помощи ЭЦВМ и использующей дескрипторный язык. Указатель табличного типа отличается большой компактностью и простотой, он легко поддается кумуляции и размножению в нужном числе экземпляров. В настоящее время для составления указателей табличного типа начинают успешно применяться счетно-перфорационные машины. Важным недостатком подобных указателей является тот факт, что работа с ними очень утомительна.

### Указатели цитированной литературы

За последнее время внимание многих ученых привлекли к себе библиографические указатели нового типа, получившие название указателей цитированной литературы (их называют также

указателями библиографических ссылок и индексами научных цитат). Наиболее полным и доступным изданием такого типа является «Science Citation Index» («Указатель цитированной литературы по точным, естественным и прикладным наукам»), выпускаемый под редакцией Ю. Гарфилда фирмой «Institute for Scientific Iinformation» (Филадельфия, США).

Широко известно, что в любой научной работе используются идеи и факты предыдущих работ в этой области. Научное использование предполагает либо новое истолкование уже имеющихся данных, либо подтверждение ими новых идей, либо дальнейшее развитие высказанных ранее мыслей, либо, наконец, опровержение правильности этих мыслей и достоверности приводимых фактов. В каждом из этих случаев ученые почти всегда ссылаются на работы своих предшественников, исследовавших до них те же, аналогичные или смежные вопросы. Эта практика возникла столетие назад и уже более полувека является обязательной нормой, служит критерием добросовестности научной публикации. Изучение библиографических ссылок в большом числе публикаций показало, что эти ссылки образуют густую сеть связей между научными документами. Группировка документов по ссылкам на одни и те же публикации позволяет выявить внутренние связи между самыми отдаленными предметами или вопросами задолго до того, как эти связи будут осознаны или изучены. Это достигается автоматически, поскольку авторы научных публикаций, посвященных самым разнообразным вопросам, могут ссылаться на один и тот же документ, который они считают исходным для своей работы.

Библиографические ссылки в статьях и книгах на использованные их авторами материалы давно служат ученым и специалистам-практикам важным средством для отыскания необходимых им источников информации. Однако путь этот долог, трудоемок и не дает исчерпывающих сведений. Библиографические ссылки становятся гораздо более эффективным средством информационного поиска, если следить за ними в обратном порядке. Обычно мы интересуемся, на какие работы ссылается автор данной публикации, но гораздо важнее знать, в каких публикациях содержатся ссылки на данную работу определенного автора. Если выявить все последующие публикации, в которых имеются ссылки на данного автора, то можно узнать широту распространения его идей, области их практического применения, а также объективно оценить значимость его работы.

На рис. 20 приведена схема использования и взаимовлияния полутора десятков работ по вопросу выявления нуклеиновых кислот, опубликованных в 1941—1960 гг. (схема составлена американским ученым Г. Алленом). Из этой схемы видно, что для получения сведений о всех этих работах по содержащимся в них библиографическим ссылкам необходимо обратиться к пристатейной

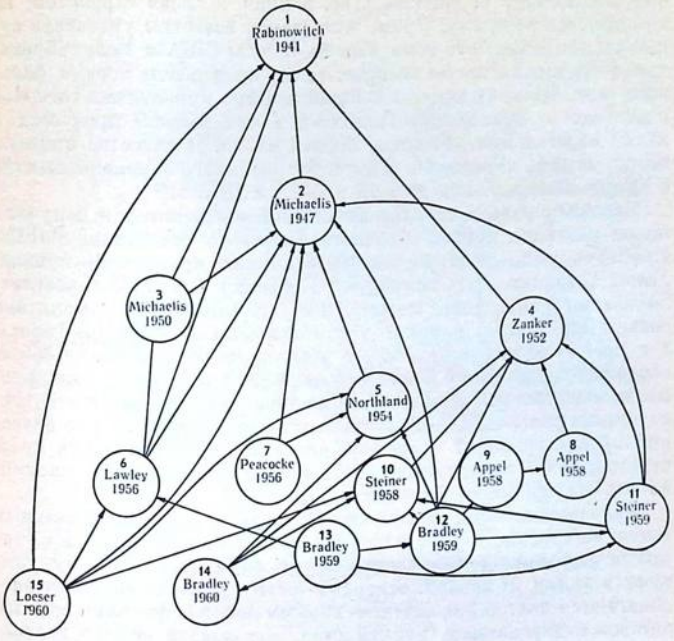



Рис. 20. Схема использования документов в последующих публикациях

библиографии по меньшей мере шести статей (6, 7, 9, 13, 14, 15), опубликованных с 1956 до 1960 г. Отдельные статьи, которые почти не упоминаются в последующих публикациях (7, 8, 9), вообще трудно выявить таким путем. С другой стороны, сведения почти обо всех этих работах (кроме 9 и 13) легко получить, если узнать, где содержатся библиографические ссылки на три ранние работы (1, 2, 4).

До последнего времени указанный принцип использовался лишь в традиционных библиографиях рецензий (известных с 1771 г.)[41], где материал всегда располагается по рецензируемым книгам, а не по рецензиям. Это понятно, так как рецензия пред-

[41] Первой библиографией рецензий была «Парижская библиография», учитывавшая по 2 журналам рецензии на книги, изданные в 1770 г. в Париже (Bibliographie parisienne. Paris, Desnos, 1771—1772).

ставляет интерес не сама по себе, а лишь в связи с работой, на которую она написана. Кроме того, давно известны указатели судебных прецедентов со ссылками на законы США и Великобритании, в судопроизводстве которых такие прецеденты играют большую роль. Поэтому юристы каждый раз при применении того или иного закона должны учитывать, в каких именно процессах и каким образом был применен данный закон. В качестве примера можно назвать «Указатель ссылок федерального законодательства о труде» Шепарда, издаваемый в США с 1873 г.[42]

Основная роль в использовании этой идеи, которая получила позже название метода «библиографического сочетания» (bibliographic coupling), а также в ее практическом применении принадлежит американскому химику Ю. Гарфилду [43]. В 1963 г. возглавляемая им фирма Institute for Scientific Information выпустила экспериментальный вариант указателя «Science Citation Index», а с 1964 г. начала издавать его регулярно. Указатель в 1965 г. содержал сведения о 3,3 млн. ссылок в 296 тыс. статьях из журналов по всем точным, естественным и прикладным наукам и в 60 тыс. патентных описаниях США. В третьем квартале 1966 г. число названий просматриваемых журналов достигло 1500. Указатель представляет собой список всех работ, которые упоминались в какой-нибудь из публикаций текущего года.

Указатель состоит из двух частей, издаваемых в виде отдельных томов: собственно «Указателя цитированной литературы» и «Указателя источников». Обе части выходят ежеквартально и кумулируются за год. В первой, основной части указателя сведения располагаются так: в алфавитном порядке даются фамилии первых авторов цитированных [44] документов, под каждой из этих заглавных фамилий приводятся (также в алфавитном порядке) фамилии первых авторов документов, содержащих ссылки на цитированный документ. Мы будем их называть «документами-источниками» или просто источниками. Сведения о работах одних и тех же авторов располагаются в прямом хронологическом порядке. Рядом с каждой фамилией указывается сокращенное название журнала, год издания, номер тома и номер первой страницы публикации. Перед годом издания приводится индекс, обозначающий тип документа-источника (статья, доклад, отчет, реферат, рецензия, обзор и т. п.).

Сложные и составные фамилии записываются как одно слово. Дефис, апостроф и другие знаки в составе фамилий опускаются. При любом числе авторов указывается лишь имя первого автора.

---

[42] Shepard's Citation. Chicago, 1873.

[43] E. Garfield. Citation indexes for science. — «Science», 1955, v. 122, № 3, стр. 108—111.

[44] Термин «цитированный документ» принят нами для обозначения документа (статьи, книги, отчета и т. д.), на который в тексте другого документа (статьи, книги, отчета и т. д.) имеется библиографическая ссылка.

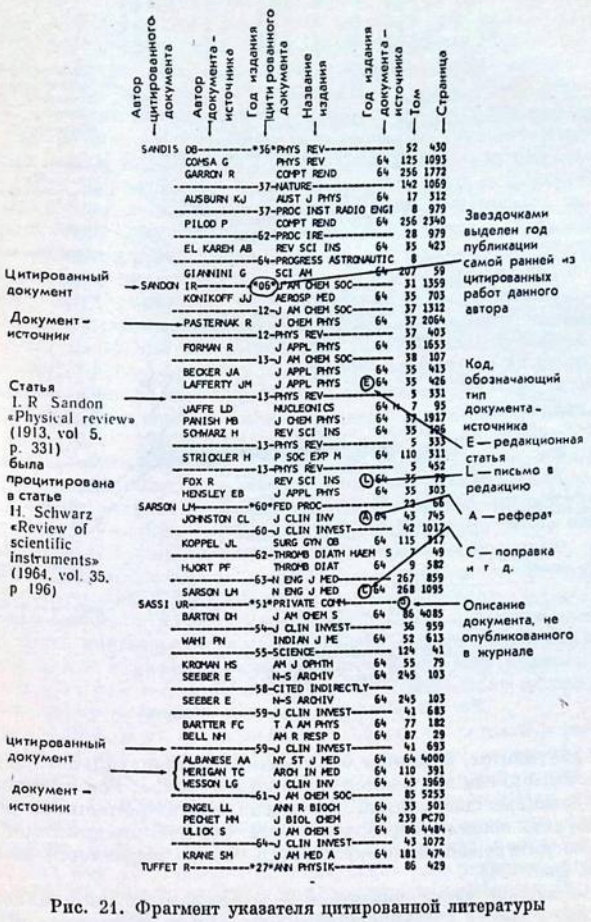



Рис. 21. Фрагмент указателя цитированной литературы

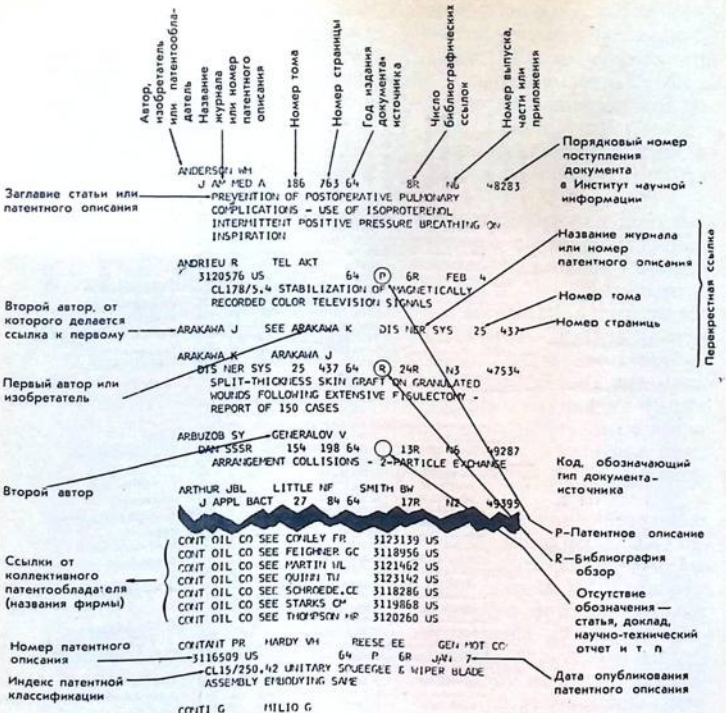



Рис. 22. Фрагмент указателя источников

Для документов, изданных без указания индивидуального автора, вместо его имени приводится слово «Anonymous». Год издания как цитированных документов, так и документов-источников дается в виде двух последних цифр этого года. Год публикации самой ранней из цитированных работ данного автора выделяется звездочками (рис. 21).

Во второй части указателя — в «Указателе источников» — описания даются в алфавитном порядке фамилий первых авторов. В отличие от первой части, где всегда указывается только фамилия первого автора, здесь приводятся фамилии всех соавторов, если их не более десяти. Характерной чертой этой части указателя

является наличие в каждом описании полного заглавия публикации. В описание включается также сокращенное название журнала, в котором опубликована статья, номер тома и страницы, год издания журнала, обозначение типа источника, число ссылок в документе, номер выпуска или части журнала (рис. 22).

«Указатель цитированной литературы» подготавливается при помощи электронных цифровых вычислительных машин. Это позволяет выпускать его в очень сжатые сроки. Объем содержащейся в указателе информации настолько велик, что обработка этой информации вручную просто невозможна. Вместе с тем, применение ЭЦВМ налагает на объем каждого описания ряд ограничений. Для каждого типа сведений отводится строго фиксированное число знаков. Например, для фамилии и инициалов автора цитированного документа — не более 18, а для тех же сведений об авторе документа-источника — всего 11.

«Указатель цитированной литературы» не предназначен для замены традиционных библиографических указателей. Он является принципиально новым средством информационного поиска. Все известные до настоящего времени библиографические указатели, предназначенные для поиска документов по их содержанию, в качестве исходного момента поиска предполагают использование специального поискового образа — классификационного индекса, предметного заголовка, набора дескрипторов и т. п. «Указатель цитированной литературы» рассчитан на использование для этой цели формальных и однозначных сведений — фамилии автора заранее установленного документа, посвященного определенному вопросу. Важнейшей особенностью этого нового указателя и является то, что, будучи построен по авторскому признаку, он дает ответы на запросы по содержанию работ. Это можно проиллюстрировать следующим примером. Допустим, что нас интересуют работы в области кибернетики и ее приложений. Достаточно знать имя известного математика Н. Винера, чтобы под этим именем найти в указателе сведения почти обо всех нужных нам работах, так как почти в каждой из них обязательно содержатся ссылки на его работы.

Эксперименты по составлению подобных указателей в разных отраслях знания проводятся в настоящее время рядом научных учреждений США: по вопросам статистики — Принстонским университетом, по вопросам атомной энергии — фирмой Itec, по вопросам психологии — Иллинойским университетом, по физике и смежным областям — Американским физическим институтом [45].

До сих пор «Указатель» использовался для решения трех типов задач: а) поиска научной информации по определенным (как правило, узкоспециальным и новым) вопросам; б) изучения раз-

[45] Е. М. Кеен. Citation indexes. «ASLIB Proceedings», 1964, v. 16, № 8, стр. 250—251.

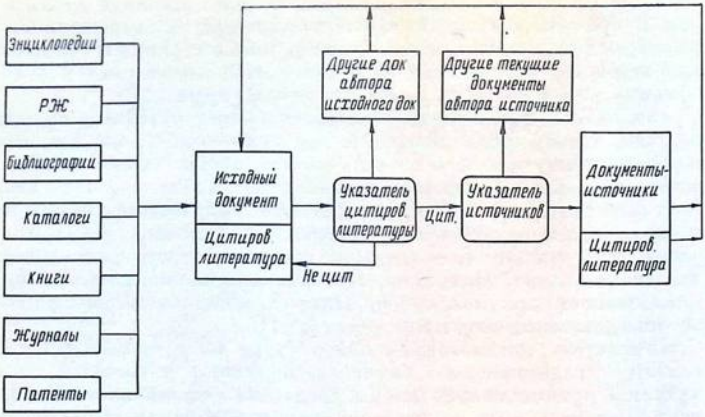



Рис. 23. Схема поиска публикаций по указателю цитированной литературы

вития определенных идей и объективной оценки их значимости; в) изучения общих закономерностей распределения и старения научных публикаций.

*Поиск* необходимых документов по «Указателю» производится следующим образом (рис. 23). На основе личного опыта или из обычных источников — энциклопедий, библиографий и т. д. — устанавливается исходный для поиска документ. Под фамилией его автора в первой части «Указателя» отыскиваются сведения о других, связанных с ним документах. Если на исходный документ никто не ссылался и он не отражен в «Указателе», то в качестве исходного документа можно взять любую имеющуюся в нем библиографическую ссылку. По второй части указателя, в которой содержатся полные описания документов-источников, определяется их полезность для ученого. Если этих документов недостаточно, то любой из наиболее ранних документов-источников становится исходным для дальнейшего поиска. Эта операция повторяется до тех пор, пока не выявляются все необходимые ученому документы, или до тех пор, пока поиск не перестанет давать новых сведений. При этом одновременно устанавливаются другие работы автора исходного (цитированного) документа, а также другие работы авторов документов-источников (помимо непосредственно разыскиваемых).

Преимущества «Указателя» как поисковой системы проявляются особенно ясно при отыскании литературы по новым узкоспециальным вопросам и направлениям. При все углубляю-

щемся взаимопроникновении наук осознание внутренних связей между отдельными предметами происходит на сравнительно поздней стадии исследований. В схемах классификации и в списках предметных заголовков эти связи проявляются еще позже в виде новых заголовков или перекрестных ссылок между прежними заголовками. В дескрипторных языках эти связи устанавливаются лишь в результате анализа эффективности работы той или иной системы в течение сравнительно большого периода времени. Между тем в «Указателях цитированной литературы» внутренние связи между самыми отдаленными предметами выявляются задолго до того, как эти связи будут осознаны исследователями. При этом совершенно устраняется субъективность мнений, неизбежная при традиционных процессах классификации и предметизации. Поэтому «Указатель цитированной литературы» является ценным источником информации по новым и особенно межотраслевым проблемам науки и техники.

Возможность *изучения развития определенных идей* при помощи «Указателя цитированной литературы» была подтверждена экспериментом, проведенным в 1964 г. Ю. Гарфилдом [46]. История одного из важнейших открытий современной биологии — механизма передачи наследственности посредством дезоксирибонуклеиновых кислот — была параллельно исследована традиционным методом и методом библиографического сочетания. Традиционный метод основывался на изучении монографии известного американского ученого и писателя Айзека Азимова «Генетический код» [47], а метод библиографического сочетания — на «Указателе цитированной литературы по генетике».

Сравнение показало, что из всех работ, выявленных обоими методами в качестве важных исторических моментов на пути к этому открытию, 65% работ являются общими. Лишь для нахождения публикаций прошлого века, когда практика ссылок на использованные работы не была еще достаточно развита, традиционный метод оказался несколько более действенным. Что же касается последних десятилетий, то «Указатель цитированной литературы» позволил установить множество связей между публикациями, упущенных в книге А. Азимова. Этот эксперимент убеждает в том, что «Указатель цитированной литературы» верно и достаточно полно отражает подлинные связи в науке. Кроме того, он подтвердил априорное предположение, что чем большее внимание исследователей привлекает та или иная публикация и, следовательно, чем чаще они на нее ссылаются, тем более серьезное научное значение она имеет.

[46] E. Garfield, I. Sher and R. Torpie. The use of citation data in writing the history of science. Philadelphia, Institute for Scientific Information, 1964.

[47] I. Asimov. The genetic code. New-York, New American Library, 1963.

одни и те же фамилии в разных случаях приводятся по-разному. Учитывая, что поиск информации в указателе осуществляется по фамилиям, можно понять, насколько это снижает его ценность. Еще одним ограничением его эффективности служит то, что все ведения в указателе приводят в алфавитном порядке фамилии авторов, указанных первыми в использованных и содержащих ссылку статьях. Вместе с тем все большее число работ создается многими авторами. Среднее число авторов на одну публикацию в настоящее время достигает двух.

Указатель цитированной литературы—сложное и дорогое издание. Его нельзя просто читать, как реферативный журнал. Он рассчитан на высококвалифицированных специалистов, и работа с ним требует больших интеллектуальных усилий. Результаты, полученные при всех видах использования такого указателя, нуждаются в глубоком осмысливании и тонкой корреляции. Однако эти результаты уникальны, так как никаким другим способом их получить нельзя. Указатели цитированной литературы имеют большую ценность для руководителей науки и ученых, определяющих основные направления научных исследований.

В нашей стране работа с указателями цитированной литературы должна развертываться в двух направлениях. Во-первых, необходимо использовать американский «Science Citation Index» для информирования советских ученых—академиков и руководителей проблемных советов и научно-исследовательских институтов АН СССР и т. д., а также для анализа структуры фронта научных исследований. Во-вторых, необходимо приступить к составлению собственного указателя цитированной литературы, так как для американского указателя обрабатывается всего 28 советских журналов, которые регулярно переводятся на английский язык и издаются в США и Великобритании. В то же время для «SCI» расписываются 559 журналов США, 170 журналов Великобритании и 49 журналов Японии. (Для сравнения укажем, что в СССР издается 1745 научно-технических журналов, в США— 6200, в Великобритании — 2200 и в Японии — 2800). Поскольку в «SCI» расписывается 1500 названий журналов, издаваемых в 39 странах мира, этот указатель может быть использован нами как система для поиска соответствующих зарубежных публикаций, а также как инструмент для изучения распространения идей советских авторов за рубежом.

Составление *советского указателя цитированной литературы* должно обеспечить прежде всего охват советских научно-технических публикаций, что даст возможность получить ответы на вопросы, на которые не отвечает американский указатель. Наш указатель цитированной литературы должен позволять отыскивать соответствующие публикации советских авторов и по этим публикациям изучать структуру исследовательского фронта советской науки. Отсюда следует, что он должен составляться на ос-

новании информации, получаемой путем обработки только советских журналов. При этом на первом этапе могут учитываться лишь ссылки на советские работы, на втором этапе — также и ссылки на зарубежные работы, что расширит базу «исходных» работ для поиска и позволит учитывать распространение идей зарубежных авторов в СССР.

### *Сигнальная информация*

В последнее время все большее распространение получает оперативное научно-информационное обслуживание, называемое *сигнальной информацией*. Оно заключается в быстром оповещении потребителей информации о появлении материалов, которые ему необходимы или могут его заинтересовать, т. е. как бы в подаче специалисту соответствующего «сигнала».

Еще сравнительно недавно потребность специалистов в быстром ознакомлении с новой литературой по интересующим их вопросам удовлетворялась в основном традиционными библиотечно-библиографическими средствами, важнейшими из которых являлись: текущие библиографические указатели; списки изданий, поступивших в библиотеку; библиографические карточки на новые книги и статьи; выставки и картотеки новых поступлений в библиотеку. Эти средства на определенном этапе играли положительную роль и до настоящего времени не потеряли своего значения.

Однако эти средства обладают некоторыми серьезными недостатками. Во-первых, сообщаемые в них библиографические сведения значительно запаздывают по сравнению с выходом первичных источников информации; во-вторых, в них, как правило, отсутствует необходимый поисковый аппарат, что приводит к необходимости сплошного просмотра всего «сигнализирующего» материала. Кроме того, выставки и картотеки новых поступлений требуют регулярного посещения библиотеки, а все остальные традиционные средства — обращения в библиотеку для получения необходимых первичных источников.

Сигнальная информация возникла в научно-информационной деятельности как реакция на эти недостатки традиционных библиографических средств, и потому одна из ее отличительных черт — это возможность для потребителя легко получить копию необходимого ему материала. Обеспечение этой возможности реализуется в настоящее время развитием служб копирования в органах научной информации и специальных библиотеках.

Сказанное выше позволяет сформулировать следующие требования к сигнальной информации. Во-первых, она должна подготавливаться в короткий срок, равный периоду накопления информации плюс несколько дней на ее обработку и в общем итоге не превышающий полутора месяцев. Во-вторых, сообщаемые сведения должны с максимально возможной полнотой охватывать

литературу по определенной отрасли знания, предмету или проблеме. В-третьих, и это, пожалуй, самое важное, сигнальная информация должна сопровождаться эффективным поисковым аппаратом, который позволял бы специалисту даже при обилии информации быстро отбирать необходимые материалы.

Дело в том, что в настоящее время по многим, даже очень узким, но актуальным вопросам появляется так много литературы, что без поискового аппарата она становится труднообозримой. В зависимости от различной организации поискового аппарата можно выделить два типа сигнальной информации: информационно-библиографические издания и избирательное (адресное) распределение информации. В информационно-библиографических изданиях сведения группируются в определенном порядке и отражаются в нескольких аспектах (по содержанию, автору, языку или времени публикаций и т. п.) при помощи соответствующих указателей.

При *избирательном (адресном) распределении информации* (selective dissemination of information) потребитель как бы доверяет поиск интересующих его сведений органу информации. Он сообщает этому органу признаки, по которым должен вестись поиск (индексы классификации, предметные заголовки, дескрипторы, имена авторов работ или цитируемых в них публикаций и т. п.). Вся литература, поступающая в орган информации, анализируется по этим признакам, и в случае их совпадения с признаками, указанными потребителем, ему регулярно посылаются только его одного интересующие сведения. Система обратной связи с потребителем позволяет информационному органу постоянно корректировать «полученную выдачу» и тем самым улучшать информационное обслуживание.

Выбор того или другого типа сигнальной информации, по-видимому, определяется спецификой информационных потребностей специалистов. При решении оригинальных научно-исследовательских проблем специалисты, как правило, предпочитают самостоятельно просматривать сведения о литературе по интересующей их проблеме и отбирать необходимые материалы. Для них более удобны сигнальные информационно-библиографические издания. Избирательное распределение информации предпочтительнее для специалистов, решающих более определенные, частные, прикладные или технологические задачи с вполне ясными границами.

Еще одной особенностью сигнальной информации является ее способность удовлетворять информационные потребности по многоотраслевым и межотраслевым проблемам. Их число в науке, особенно в технике, увеличивается и привлекает все бóльшие массы ученых и инженеров. Однако реферативные журналы продолжают издаваться по отраслевому признаку и, очевидно, будут и в дальнейшем сохранять этот принцип. Издание реферативных журналов

по отдельным межотраслевым проблемам слишком сложно и дорого, тем более что число этих проблем достаточно велико и имеет тенденцию к увеличению. Вместо реферативных журналов по много- и межотраслевым проблемам целесообразно выпускать издания сигнальной информации, которые значительно дешевле, компактнее и потому допускают дублирование информации в гораздо бо́льших размерах.

Ниже приводится краткая характеристика основных типов сигнальной информации с примерами наиболее известных зарубежных и отечественных информационных изданий и служб.

*Систематические* или *предметные библиографические указатели* являются самым распространенным видом подобных изданий. В этих указателях заглавия или полные библиографические описания публикаций приводятся под заголовками схем классификации или предметными заголовками с повторным отражением в авторских или предметных вспомогательных указателях. Примерами могут служить описанный выше указатель медицинской литературы «Index Medicus» (США), указатель литературы по физике «Current Papers in Phisics» (Великобритания), а также систематизированный библиографический указатель ВИНИТИ по физике элементарных частиц и теории полей (СИ «Частицы и поля»).

СИ «Частицы и поля» выходит с 1967 г. ежемесячно и состоит из библиографического и авторского указателей.

Описания статей из нескольких десятков отечественных и зарубежных журналов публикуются на языке оригинала и систематизируются по специальной схеме классификации, содержащей 250 заголовков. Указатель можно использовать для заказа в ВИНИТИ копий или переводов оригинальных статей.

*Пермутационные указатели*, описанные выше, часто входят как основная составная часть в издания сигнальной информации. Наряду с широко известным изданием «Chemical Titles», можно назвать двухнедельный бюллетень «B.A.S.I.C.» (Biological Abstracts Subject Index in Context — Указатель содержания рефератов по биологии в контексте), издаваемый с 1962 г. редакцией реферативного журнала «Biological Abstracts». Этот указатель публикуется в каждом выпуске реферативного журнала вместо прежнего предметного указателя и охватывает ежегодно около 100 тыс. рефератов.

С 1967 г. ВИНИТИ приступило к публикации подобного издания сигнальной информации по автоматике и радиоэлектронике в трех выпусках: «Автоматика, телемеханика и вычислительная техника», «Электроника и ее применение» и «Радиотехника». Каждый выпуск этого издания выходит два раза в месяц и состоит из пермутационного, библиографического и авторского указателей. Пермутационный и авторский указатели служат для поиска по содержанию и фамилиям авторов интересующих потребителей

публикаций, библиографический указатель содержит полные описания статей на языке оригинала, расположенные в алфавитном порядке названий журналов (в каждом выпуске отражается содержание нескольких десятков названий отечественных и зарубежных журналов).

*Указатели содержания журналов* составляются из оглавлений только что вышедших в свет номеров основных журналов по определенным отраслям знаний. Наиболее известное из подобных изданий — «Current Contents» — с 1958 г. выпускается фирмой «Institute for Scientific Information» (США). В настоящее время это еженедельное издание, снабженное указателем авторов и их адресов, выходит в трех сериях. «Current Contents, Physical Sciences» помещает сведения о содержании свыше 650 названий журналов (2 тыс. статей в неделю) по астронавтике, физике, атомной энергии, математике, приборостроению, электронике, вычислительной технике, комплексу наук о Земле, металлургии и смежным дисциплинам. «Current Contents, Life Sciences» помещает сведения о содержании свыше 780 названий журналов по фармакологии, физиологии, медицине, биохимии, биологии, генетике, ботанике, экологии и смежным дисциплинам. «Current Contents, Chemical Sciences» издается с 1967 г. и помещает сведения более чем о 50 тыс. статей в год из ведущих журналов по химии и химической технологии. В дополнение к обычным указателям эта серия в каждом номере имеет предметный указатель, а заглавия статей сопровождаются структурными формулами рассматриваемых в них химических соединений.

К этому же типу изданий может быть причислен выпускаемый той же фирмой еженедельный указатель новых химических соединений «Index Chemicus». Указатель сообщает о появлении какого-либо соединения через один-два месяца после того, как сведения о нем были опубликованы в мировой специальной литературе. Записи расположены в алфавитном порядке названий журналов, в которых опубликованы соответствующие статьи. Каждая запись содержит библиографическое описание статьи, ее краткий реферат, структурные формулы новых соединений, записи реакций и некоторые дополнительные сведения о механизме реакций и физико-химических свойствах соединений. Большинство записей в указателе предварительно просматривается авторами статей. Все сведения ежемесячно отражаются в формульных, авторских и предметных указателях, а затем кумулируются за полгода и за год. За период с 1960 по 1966 г. в издании были помещены сведения о 700 тыс. новых соединений. Эти сведения собраны в специальном ежегодном кумулятивном издании «Encyclopaedia Chimica Internationalis» (23 тома) со специальными указателями.

С 1967 г. в ВИНИТИ выпускается сигнальное издание подобного типа — двухнедельный бюллетень СИ «Электрохимия и электрохимические производства», включающий библиографи-

ческие описания всех отечественных и зарубежных материалов, которые в дальнейшем будут прореферированы в двух разделах РЖ ВИНИТИ «Химия»: «Электрохимия» и «Химические источники тока. Электрохимические производства. Электроосаждение». Описания статей расположены в алфавитном порядке названий журналов, а описания авторских свидетельств и патентов — в конце каждого выпуска: отечественные в порядке номеров, а зарубежные — в алфавитном порядке названий стран. Каждый выпуск содержит 200—300 описаний.

Критериями выбора определенного вида сигнальной информации для конкретных отраслей науки и техники могут служить: информативность заглавий статей в данной отрасли и сложившиеся в ней традиции, подготовленность читателей к восприятию сигнальной информации того или иного вида. Наибольшей подготовленности читателей требует использование пермутационных указателей. Однако они имеют то преимущество, что их подготовка может быть полностью автоматизирована.

В заключение необходимо сказать несколько слов об избирательном (адресном) распределении информации. Примером такого обслуживания может служить система американской фирмы «Institut for Scientific Information» под названием ASCA (Automatic Subject Citation Alert — Автоматическое оповещение по содержанию и цитированной литературе). Эта система еженедельно предоставляет каждому индивидуальному подписчику сведения о всех публикациях, которые содержатся в получаемой фирмой литературе и соответствуют признакам, указанным каждым потребителем: ключевым словам заглавий или именам авторов цитированных работ.

В СССР пока нет опыта широкого использования этого вида сигнальной информации. В небольших масштабах она успешно развивается в низшем звене информационной сети: информационных органах научно-исследовательских институтов, конструкторских бюро, заводов, а также в некоторых учебных заведениях, например, в Московском государственном университете им. М. В. Ломоносова на Кафедре научной информации. Особенно важно развивать избирательное распределение информации по неопубликованным документам, так как без «принудительного» распределения непубликуемая информация часто оказывается недостаточно использованной.

Избирательное распределение информации в сочетании с оперативным копированием документов по заказам потребителей может стать одним из важнейших видов информационного обслуживания специалистов, если оно будет основано на механизированных отраслевых информационно-поисковых системах крупного масштаба.

Глава 3

# АНАЛИТИКО-СИНТЕТИЧЕСКАЯ ПЕРЕРАБОТКА ДОКУМЕНТОВ

Аналитико-синтетическая переработка документов состоит в представлении каждого отдельного документа или определенной совокупности документов в таком виде, который максимально отвечает той или иной задаче научно-информационной деятельности. Основными видами аналитико-синтетической переработки являются библиографическое описание документов, их классификация (индексирование), аннотирование, реферирование, перевод с одного языка на другой, а также составление обзоров. Каждому виду такой переработки свойственно определенное соотношение анализа и синтеза.

При использовании научных документов прежде всего возникает необходимость называть их, ссылаться на них, производить идентификацию документов (т. е. устанавливать тождество нескольких копий одного документа без их непосредственного сличения), а также составлять перечни документов. Для этого используется библиографическое описание, которое содержит необходимый и достаточный набор определяющих документ признаков. Для раскрытия содержания научного документа путем его отнесения к некоторому классу документов, посвященных какому-либо одному вопросу или относящихся к определенной отрасли знания, служит классификация (систематизация). Такая принадлежность документа к некоторому классу документов, родственных ему по содержанию, может быть обозначена путем приписывания (присвоения) этому документу соответствующего индекса (кода). Если процесс классификации документов включает присвоение им соответствующих индексов (кодов), то такой процесс называется индексированием. Библиографическое описание и индекс документа недостаточно раскрывают его содержание. Более полная характеристика содержания документа может быть дана в результате его аннотирования и реферирования.

Перечисленные виды аналитико-синтетической переработки документов совершенствовались в течение столетий в связи с раз-

витием книгоиздательского, книготоргового и библиотечного дела, библиографии, прикладной лингвистики и организации научного труда. В настоящее время такие виды переработки, как библиографическое описание, классификация, перевод с одного языка на другой, оформились в самостоятельные научные дисциплины, имеющие свою теорию и методику. Однако только с возникновением информатики удалось выявить их внутреннюю взаимосвязь, принадлежность этих видов аналитико-синтетической переработки к одному классу операций, а также поставить вопрос об их механизации и автоматизации.

Ниже кратко излагаются история, теория и методика основных видов аналитико-синтетической переработки документов, а также даются рекомендации по практической работе с документами.

## БИБЛИОГРАФИЧЕСКОЕ ОПИСАНИЕ

Термин «библиографическое описание» многозначен. Прежде всего он означает некоторый необходимый минимум сведений о документе, выраженных словами и расположенных в строго установленном порядке. В этом смысле описание служит основой для составления библиотечных каталогов, библиографических картотек и указателей, входит как неотъемлемая часть во вторичные документы — аннотации, рефераты, обзоры и т. п. Этим термином обозначают также один из видов аналитико-синтетической переработки документов, а также специальную библиотечно-библиографическую дисциплину.

Библиографическое описание — чрезвычайно ответственный и трудоемкий процесс. Он требует хорошего знания всех видов первичных и вторичных документов, основ полиграфического производства, умения определять содержание, форму и назначение документа и находить данные, необходимые для его правильного и полного отражения. При описании документов для библиотечных каталогов и картотек используются главным образом сведения, содержащиеся в самом документе. Это называется описанием «de visu», т. е. «имея документ перед глазами». В случае необходимости каталогизатор добавляет недостающие сведения, выявляет в документе библиографические сведения о других документах, определяет дефекты документа (если они имеются) и выясняет, относятся они ко всему изданию или только к описываемому экземпляру. Описание для других вторичных документов, кроме библиотечных каталогов, может производиться не по самому документу, а на основе каталогизационного описания. О трудоемкости библиографического описания документов для библиотечных каталогов свидетельствует то, что один человек за рабочий день может описать не более 50—70 книг, а в особых случаях (книги на иностранных языках, старопечатные издания) — не более 20—30 названий.

графического порядка для Санкт-Петербургской императорской библиотеки» ее будущий директор А. Н. Оленин опубликовал первую отечественную инструкцию по каталогизации. В каталогизационных работах этой библиотеки принимали участие выдающиеся деятели русской культуры А. Х. Востоков, М. Н. Загоскин, И. А. Крылов, В. С. Сопиков, В. Г. Анастасевич, В. И. Собольщиков, В. В. Стасов. Высказанные в ходе этих работ идеи в дальнейшем были разработаны западными библиотекарями и библиографами и легли в основу современных принципов каталогизации. В частности, в каталогах петербургской Публичной библиотеки первой четверти XIX в. некоторые издания были отражены под названием опубликовавших их учреждений (под коллективным автором).

Дальнейшее развитие теория книгоописания получила в работах, проводившихся в связи с составлением каталога библиотеки Британского музея. Особые заслуги принадлежат директору этой библиотеки А. Паницци, который в 1841 г. опубликовал «Правила составления каталога печатных книг Британского музея» [1] — наиболее детальную инструкцию того времени, в которой принципы описания поставлены в зависимость от сформулированных в предыдущей главе задач алфавитного каталога. В этой инструкции впервые начата теоретическая разработка принципа описания некоторых изданий под коллективным автором, который до этого применялся только эпизодически. Еще более четко сформулировал новые идеи известный американский библиотековед Ч. Кеттер (1837—1803 г.). «Теперь все правила, — писал он, — могут быть сведены к двум основным принципам: первый состоит в том, что книги должны описываться под именем их автора или редактора (в случае сборников) или же под названием учреждения, ответственного за их издание, а второй принцип состоит в том, что если авторы неизвестны, то в качестве заголовка выбирается первое слово заглавия, не являющееся артиклем или предлогом» [2].

В начале XX в. была создана так называемая англо-американская система книгоописания, которая в дальнейшем получила широкое распространение. В это же время сложилась и «прусская» система книгоописания, в которой нашли отражение европейские каталогизационные традиции. Создателем «прусской» системы книгоописания, ныне утратившей свое значение, был директор университетской библиотеки во Вроцлаве К. Дзяцко (1842—1903 гг.). Его взгляды на каталогизацию отличались двумя характерными чертами: ограниченным истолкованием понятия

---

[1] A. Panizzi. Rules for the compilations of the catalogue. В кн.: «Catalogue of printed books in the British Museum», v. 1. London, 1841, стр. V—IX.

[2] Ch. A. Cutter. The new catalogue of Harward College Library. «North American Review», 1869, № 108, стр. 102—104.

«автор» и своеобразной методикой описания анонимных произведений. Под автором он понимал лишь лицо, непосредственно написавшее книгу, и поэтому отрицал возможность вынесения в авторский заголовок имен редакторов, издателей и других лиц, принимавших участие в создании книги. Кроме того, К. Даицко отвергал принцип описания под названием учреждений (под коллективным автором). Анонимные произведения он предлагал описывать под первым грамматически независимым существительным в именительном падеже (Substantivum regens). Рекомендованные им принципы не были прогрессивными, что и обусловило недостатки основанной на них «прусской» системы каталогизации.

В отечественной теории книгоописания переломным моментом явилась Октябрьская социалистическая революция. С самого начала каталогизационных работ в СССР был взят курс на коллективное создание государственного свода правил книгоописания. Широко обсуждая возможности использования зарубежного опыта, советские библиотекари учитывали необходимость опираться на национальные традиции, специфику русского языка и особенности советской печатной продукции. Эта задача была решена Московской каталогизационной комиссией, работавшей с 1922 г. под председательством проф. А. И. Калишевского, которого затем сменил Г. И. Иванов. Одним из наиболее активных ее членов на протяжении трех с половиной десятков лет был известный библиотековед Л. В. Трофимов. Работа комиссии по выработке первой советской государственной инструкции протекала в условиях исключительного интереса советских библиотековедов и библиографов к вопросам книгоописания и, в частности, к принципу описания под коллективным автором. Это было вызвано тем, что в условиях советского государства официальные издания учреждений и организаций приобретали все большее значение. Важную роль в разработке методики описания под коллективным автором сыграла Всесоюзная книжная палата, в частности, один из ее ведущих работников проф. Е. И. Шамурин, многие идеи которого были реализованы в современных сводах правил книгоописания. Уже в первой государственной инструкции, выходившей с 1926 по 1932 г., были закреплены методы описания, теоретически обоснованные советскими библиографами и библиотековедами в предшествующие годы: описание всех произведений одного индивидуального автора под единообразной формой его имени, описание под коллективным автором официальных и ведомственных изданий, а также описание анонимных произведений под первым словом заглавия. Тем самым советская каталогизация решительно порвала со старыми европейскими традициями, господствовавшими в дореволюционных библиотеках.

С выходом «Инструкции» разработка методики описания не только не приостановилась, но стала еще более интенсивной. В «Инструкции» были определены лишь общие принципы, требо-

вавшие дальнейшего уточнения и детализации. В результате многолетней работы в настоящее время в СССР создано несколько сводов правил описания произведений печати, основанных на одинаковых принципах и отличающихся лишь полнотой и подробностью правил. Они должны способствовать составлению библиотечных каталогов и библиографических указателей всех типов, а также могут быть использованы при разработке методов машинного библиографического описания научных публикаций.

### *Общие требования и основные принципы*

К библиографическому описанию документов предъявляются некоторые общие требования, которые не зависят ни от вида документов, ни от назначения описания. Кратко эти требования можно сформулировать следующим образом: 1) описание должно точно и полно отражать содержание и форму документа; 2) описание документа должно обеспечивать возможность безошибочной и быстрой идентификации (установления тождественности) каждого его издания; 3) разные описания одного и того же документа должны обязательно содержать все основные элементы; такие описания могут отличаться порядком расположения этих элементов и подробностью приводимых в них сведений.

Смысл этих требований может быть раскрыт путем рассмотрения специфических различий в описаниях, составляемых для библиотечных каталогов и библиографических указателей. Библиографические указатели по сравнению с каталогами имеют гораздо более узкое и точное назначение, например повышение квалификации рабочих определенной профессии или информацию ученых о новых исследованиях. При просмотре карточных каталогов каждое описание, занимающее отдельную карточку, воспринимается самостоятельно. Сведения на карточке располагаются поразному в зависимости от вида описания. В библиографическом же указателе, имеющем форму книги или списка, все описания, независимо от их вида, располагаются на странице по одной общей схеме и воспринимаются в комплексе. Это дает возможность вынести в общий заголовок и не повторять в каждом описании такие совпадающие сведения, как фамилия автора, год издания или заглавие серии. Вот почему описание для библиографических указателей может несколько отличаться от описания для библиотечного каталога. Оно зависит и от назначения библиографии, и от охвата материала по времени и характеру изданий, и от многих других факторов (это, разумеется, не относится к тем случаям, когда карточки для каталога и страницы библиографии печатаются с одной формы).

Вместе с тем нельзя забывать, что потребитель информации часто узнает о том или ином документе именно из библиографиче-

ских источников, после чего он обращается за этим документом в библиотеку, разыскивая его по библиотечному каталогу. Следовательно, если документ будет по-разному описан в библиографии и каталоге, то специалист может его не найти. В библиографическом указателе описание сборника статей разных авторов часто дается под фамилией его составителя. В каталоге же этот сборник будет описан под заглавием. Потребитель информации может не найти этого сборника в каталоге и решить, что его нет в справочно-информационном фонде.

Проблема взаимосвязи каталогизационного и библиографического описания решается таким образом, чтобы для подавляющего большинства видов изданий основное описание в каталоге и библиографическая запись в указателе начиналась с одних и тех же слов. Когда различные требования к описанию не позволяют достигнуть необходимого единообразия, в алфавитных каталогах обязательно предусматривается добавочное описание под теми сведениями, с которых может начинаться описание документа в библиографических указателях.

*Основные принципы книгоописания*, сложившиеся к настоящему времени, определяют выбор сведений, которые содержат самые важные для данного документа признаки и поэтому должны стоять в самом начале описания. Они также определяют место описания данного документа в алфавитном ряду других описаний. Таких признаков три: а) имя индивидуального автора документа; б) название учреждения или организации, ответственных за издание документа, и в) заглавие документа.

Принцип авторского описания состоит в том, что имя автора, в каком бы месте документа оно ни было указано, обязательно выносится в заголовок основного описания. Этот принцип является наиболее распространенным. Некоторые документы создаются более чем одним автором, а другие выпускаются без указания автора. В связи с этим возникает вопрос, кого следует считать автором произведения и какие именно документы следует описывать под индивидуальным автором. Иногда авторы указываются в различных изданиях их работ под разными именами или под несколькими именами. В связи с этим возникает другая группа вопросов: в какой форме следует выносить в заголовок имя автора, чтобы заголовки основных описаний произведений данного автора были составлены единообразно.

Принцип описания под коллективным автором, получивший теперь почти всеобщее распространение, позволяет наиболее точно отразить во всех вторичных научных документах исключительно важный вид произведений печати — официальные издания. Это объясняется тем, что наиболее важной библиографической характеристикой официальных изданий является название опубликовавшего их учреждения или организации, т. е. «коллектива». При этом название коллектива служит наилучшим признаком не только для

тех изданий, заглавия которых не индивидуализированы и начинаются с типовых слов («отчет», «труды», «материалы» и т. п.), но и для многих изданий, имеющих характерные заглавия. В алфавитных каталогах, библиографических указателях и других вторичных изданиях с алфавитным расположением материала применение этого принципа обеспечивает группировку описаний всех публикаций, в которых в качестве автора выступает один и тот же коллектив. Это имеет большое значение для выявления научной, литературной и практической деятельности учреждений и организаций и оказывает существенную помощь при поиске официальных и ведомственных изданий.

Принцип описания под заглавием является одним из наиболее древних и до сих пор доминирует в практике книгоописания многих восточных стран. При использовании этого принципа возникает два вопроса: а) какие издания целесообразно описывать под заглавием и б) в какой форме следует приводить заглавие. Обычно под заглавием описываются такие документы, которые запоминаются и запрашиваются именно по заглавию. Наиболее характерными примерами таких изданий могут служить произведения четырех и более авторов, сборники произведений разных авторов, периодические издания и стандарты. Что же касается формы заглавия, то сейчас почти повсеместно применяется порядок слов заглавия, данный в самом документе. До последнего времени в ряде стран Центральной Европы применялся так называемый грамматический порядок, при котором на первое место выдвигалось слово, несущее основную смысловую нагрузку (вместе с группой примыкающих к нему слов). Этим словом служило первое независимое существительное в заглавии, не являющееся определением или обстоятельством. В качестве примера можно привести название данной главы, в котором цифрами обозначен грамматический порядок слов:

2 1 3

«Аналитико-синтетическая переработка документов»

Применение грамматического порядка слов в заглавии (инверсия) — методический прием, разработанный в ходе многолетней библиотечно-библиографической практики. К инверсии (перестановке слов в заглавии) неоднократно прибегали отечественные и зарубежные библиографы еще в прошлом веке. Давая известные преимущества при группировке описаний анонимных книг с заглавиями, начинающимися с несущественных определений, этот прием затрудняет идентификацию документов и нередко превращает библиографическое описание в сложнейший ребус. Поэтому в настоящее время инверсия слов в заглавии почти не применяется при описании документов.

### *Международные принципы*

Унификация принципов библиографического описания имеет большое значение для развития контактов между информационными центрами, библиографическими учреждениями и библиотеками разных стран. В 1961 г. в Париже проходила Международная конференция по принципам каталогизации. На этой конференции почти единодушно была принята резолюция, в которой содержится изложение международных принципов книгоописания. Активное участие в работе конференции принимала советская делегация, доклад которой «Границы применения описания под коллективным автором» был одобрен большинством участников конференции. Почти все рекомендации конференции по существу подтверждают принципы, положенные в основу советских «Единых правил».

Основная трудность достижения международного соглашения о принципах книгоописания состояла в том, что две главные зарубежные системы книгоописания — англо-американская и прусская, сложившиеся еще в начале нашего века, расходились по коренным вопросам каталогизации. Сильными сторонами англо-американской системы нужно признать применение описания под коллективным автором и простоту алфавитного расположения описаний под заглавием. Благодаря этому она оказала заметное влияние на каталогизационную практику многих стран. Однако неумеренное расширение круга изданий, описываемых под коллективным автором, и сложность правил, определяющих форму названий учреждений в заголовке описания, сильно снижали эффективность англо-американской системы. Именно поэтому она встречала противодействие со стороны библиотекарей и библиографов многих стран Европы и в первую очередь немецких библиотекарей, которые до последнего времени предпочитали придерживаться старых европейских традиций, сформулированных в прусской инструкции.

С другой стороны, основные положения прусской системы — отказ от описания под коллективным автором и применение грамматического порядка слов в заглавии — давно уже подвергались критике со стороны библиотековедов и библиографов ГДР, ФРГ и ряда других стран — Австрии, Дании, Норвегии, Польши, Чехословакии, Швейцарии, Швеции. Главное достижение конференции в Париже как раз и заключалось в том, что это противоречие было устранено. Достигнута также договоренность и по многим другим важным вопросам.

Сформулированные этой конференцией принципы предназначены специально для каталогов, раскрывающих фонды крупных универсальных библиотек. Но эти принципы рекомендуется применять также во вторичных изданиях с алфавитным расположением материала. Произведения одного автора должны описываться под

наиболее известным и полным именем этого автора. Основное описание [3] под коллективным автором рекомендуется применять в тех случаях, «когда произведение по своему характеру является выражением коллективной мысли или деятельности коллектива (например, официальные отчеты, правила и распоряжения, манифесты, программы и протоколы)... или когда формулировка заглавия или данные титульного листа в совокупности с характером произведения ясно показывают, что коллектив в целом несет ответственность за содержание произведения (например, периодические издания, содержащие некоторые данные о деятельности коллектива, названия которых состоят из типовых слов: «бюллетень», «труды» и т. п.» [4].

Согласно международным принципам каталогизации, основное описание под заглавием получают: 1) произведения, авторы которых не установлены; 2) произведения более чем трех авторов, ни один из которых не является основным автором; 3) сборники произведений или частей произведений разных авторов, изданные под общим заглавием; 4) произведения печати (включая продолжающиеся и периодические издания), более известные по заглавию, чем по имени автора.

### *Методика библиографического описания*

Для удовлетворения общих требований, предъявляемых к описанию документов, а также для реализации основных принципов книгоописания, необходимо руководствоваться определенными правилами. Эти правила помогают одинаково описывать одни и те же произведения печати, независимо от того, где, кем и для какой цели составляется описание. Совокупность правил описания называется сводом правил или каталогизационной инструкцией. В каждой стране имеется несколько таких инструкций, которые служат библиотекарям и библиографам руководством в практической работе по описанию. Работники органов научной информации, у которых нет своих каталогизационных инструкций, приспосабливают для своих надобностей существующие правила библиографического описания.

В СССР создана система каталогизационных инструкций, которые называются «Едиными правилами» или просто «Правилами». В крупных универсальных и специальных библиотеках приме-

---

[3] Основным описанием называется описание, наиболее полно характеризующее документ и используемое для его отражения во всех каталогах и картотеках.

[4] Международная конференция по принципам каталогизации. Париж, 9—18.X 1961 г. Предварительный официальный отчет. Резолюция 1. Изложение принципов. «Бюлл. ЮНЕСКО для библиотек», 1962, 16, 2, стр. 60; International Conference on Cataloguing Principles. Paris. 1961. Report. London, (IFLA), 1963.

няются «Единые правила описания произведений печати для библиотечных каталогов» (М.—Л., 1949—1966), состоящие из шести самостоятельных частей и трех примыкающих к ним инструкций. В этот свод входят правила описания книг, периодических, картографических и нотных изданий, графики и специальных видов технических изданий, правила организации алфавитного каталога, инструктивные указания по групповой обработке, по описанию книг на языках стран зарубежного Востока, а также по описанию старопечатных книг [5].

Для небольших библиотек, в фондах которых нет многих из перечисленных видов произведений печати и которые поэтому не нуждаются в столь подробных рекомендациях, созданы «Единые правила описания произведений печати и организации алфавитного каталога для небольших библиотек» (М., 1963). В этой небольшой по объему, но достаточно полной инструкции содержатся все основные правила, применяемые при описании книг, периодических изданий, глав, разделов и статей, а также при организации алфавитного каталога. Она может быть взята за основу для изучения советской методики книгоописания.

При описании литературы для библиографических указателей следует руководствоваться специальными «Правилами библиографического описания произведений печати» (М., 1964), в которых изложены методы отражения в библиографии статей, рецензий, а также оформления сведений о литературе в подстрочных примечаниях небиблиографических изданий.

Необходимо также упомянуть о существовании упрощенных «Правил описания произведений печати для сельских библиотек» (М., 1964). Таким образом, советская система каталогизационных инструкций состоит из четырех сводов правил, которые систематически перерабатываются и выходят в новых изданиях. Необходимость в подобной работе, осуществляемой Межведомственной библиотечной каталогизационной комиссией, вызывается тем,

---

[5] Единые правила описания произведений печати для библиотечных каталогов. Ч. 1—6. М.—Л., 1949—1967.

Ч. 1, вып. 1. Описание книг. М., 1949; изд. 2. М., 1959; дополнение. М., 1967.

Ч. 1. вып. 2. Организация алфавитного каталога книг. М., 1952; изд. 2, испр. и доп. М., 1961.

Ч. 2. Описание периодических изданий. Л., 1954.

Ч. 3. Описание картографических изданий. Л., 1950.

Ч. 4. Описание нотных изданий. М., 1952; Описание нот и музыкальных записей, Изд. 2, испр. и доп. М., 1963.

Ч. 5. Описание печатной графики. Л., 1958; изд. 2, испр. и доп. М., 1967.

Ч. 6. Специальные виды технических изданий. М., 1957.

Групповая обработка произведений печати. Инструктивно-методические указания. М., 1958.

Инструкция по описанию книг на языках стран зарубежного Востока. М., 1960.

Описание иностранных книг второй половины XVI и XVII в. М., 1967.

что появляются новые типы и виды изданий, изменяется издательское оформление печатной продукции, а также совершенствуется теория и методика библиографического описания.

Научная разработка вопросов книгоописания ведется в нашей стране многими библиотечно-библиографическими учреждениями. В марте-апреле 1965 г. проходила первая в СССР научная конференция по каталогизации, которая была созвана в Москве Государственной библиотекой СССР им. В. И. Ленина и Междубиблиотечной каталогизационной комиссией. В ее работе приняло участие свыше 600 человек от 170 учреждений страны. О содержании обсужденных вопросов дает представление перечень докладов, прочитанных на пленарных заседаниях: «Состояние и перспективы развития каталогизации в СССР», «Единство принципов и специфика каталогизации произведений печати для разных типов библиотек СССР», «К вопросу о сближении каталогизационного и библиографического описания», «Механизация библиотечно-библиографических процессов и вопросы каталогизации», «Методика описания на печатных карточках централизованной каталогизации», «Современное состояние каталогизации за рубежом». Методика описания отдельных видов произведений печати была детально обсуждена на секционных заседаниях [6].

Методика описания делится на общую, рассматривающую правила описания всех произведений печати, и частную, касающуюся правил описания отдельных типов и видов изданий. В общую методику входят основные положения об источниках, видах и элементах описания, его языке и орфографии, а также правила составления отдельных элементов описания. Источником описания считается документ в целом, однако для описания в первую очередь используются сведения с титульного листа или первой страницы документа.

Описание содержит следующие элементы:

1. Сведения об авторах (заголовок описания).
2. Заглавие.
3. Подзаголовочные данные — сведения, поясняющие и уточняющие тему и содержание документа, его характер и назначение; сведения об утверждении или введении в действие официальных документов; сведения о лицах, участвовавших в составлении документа и его подготовке к изданию; сведения о повторности или официальном характере издания.
4. Выходные данные — сведения о месте издания, наименовании издательства и годе издания.
5. Сведения о числе страниц и наличии иллюстративного материала (количественная характеристика).

---

[6] Е. А. Новикова. Научная конференция по каталогизации в СССР. «Бюлл. ЮНЕСКО для библиотек», 1967, т. 21, № 1, стр. 24—31.

Образец титульного листа

Государственный Комитет
по координации
научно-исследовательских
работ СССР

Академия наук
Союза Советских
Социалистических
Республик

Всесоюзный институт научной и технической информации

*Надзаголовочные данные*

ПОСОБИЕ
ПО РЕФЕРИРОВАНИЮ И РЕДАКТИРОВАНИЮ
НАУЧНО-ТЕХНИЧЕСКОЙ ЛИТЕРАТУРЫ

*Заглавие*

*(Справочник для редакторов, референтов и корректоров)*

*Подзаголовок*

*К. т. н. В. П. Горькова, Т. В. Муравинский*

*Сведения об авторах*

Москва 1964

*Выходные данные*

6. Надзаголовочные данные — наименование учреждения или организации, ответственной за издание документа, заглавие серии, в которую входит документ, и его порядковый номер в этой серии.

7. Примечания — сведения об авторах, приведенные в документе (если они расходятся со сведениями, принятыми для описания, или если документ написан более чем двумя авторами); сведения о разночтении в заглавии; сведения о наличии в документе библиографии; сведения о дефектах в издании.

Образец описания на каталожной карточке

Горькова В. И. и Муранивский Т. В.

Пособие по реферированию и редактированию научно-технической литературы. (Справочник для редакторов, референтов и корректоров). М., 1964.

268 стр. (Всесоюз. ин-т научной и техн. информации).

Образец описания в библиографическом указателе

**Горькова В. И. и Муранивский Т. В.** Пособие по реферированию и редактированию научно-технической литературы. (Справочник для редакторов, референтов и корректоров). М., 1964, 268 стр. (Всесоюз. ин-т научной и техн. информации).

Из приведенного примера видно, что в отличие от описания для библиографического указателя, где все элементы следуют непосредственно один за другим, на каталожной карточке соблюдается определенный порядок расположения элементов описания. Это позволяет выделять каждый элемент (кроме подзаголовочных данных) и отличать один вид описания от другого, не читая текста самого описания. Благодаря этому экономится много времени при работе с каталогами, когда приходится просматривать сотни и тысячи карточек по разным элементам. В библиографическом указателе обычно применяется только одно описание; для поисков сведений о данном документе по другим признакам служат вспомогательные указатели. В алфавитном каталоге, наряду с основным описанием, для этой цели применяются добавочные описания. Добавочное описание для алфавитного каталога к приведенному образцу основного описания будет иметь следующий вид.

В частной методике рассматриваются правила описания документов, изданных с указанием или без указания авторов или составителей; правила описания многотомных и периодических изданий, официальных и ведомственных изданий; правила анали-

Образец добавочного описания

| | Муранивский Т. В. |
|---|---|
| Горькова В. И. и Муранивский Т. В. Пособие по реферированию и редактированию научно-технической литературы. М., 1964. | |

тического описания и правила описания специальных видов технической литературы.

Документы одного, двух и трех авторов обычно описываются под их именами. Следует иметь в виду, что в некоторых документах — словарях и справочниках, учебных и методических пособиях, руководствах, материалах по технической информации и обмену опытом — авторы по традиции называются составителями.

Поэтому такие документы описываются под составителями как под авторами. В других видах документов — сборниках работ разных авторов, ведомственных материалах и инструкциях, учебных программах, статистических справочниках — составители лишь подбирают результаты работы других лиц и поэтому не являются авторами. Эти виды документов описываются под их заглавиями. Под заглавием описываются также документы, принадлежащие четырем и более авторам, причем имена авторов указываются в примечании к описанию (четыре — в документах четырех авторов и три — в документах пяти и более авторов).

Образец описания под заглавием на каталожной карточке

| | Справочник автомобильного механика. Под общ. ред. |
|---|---|
| Л. Л. Афанасьева. Изд. 3, переработ. и доп. М., Машгиз, 1959. 892 стр. с илл.; 1 л. черт. Перед. загл. авт.: Л. Л. Афанасьев, В. А. Илларионов, Н. Э. Струве и К. С. Шестопалов. | |

Образец описания под заглавием в библиографическом указателе

**Справочник** автомобильного механика. Под. общ. ред. Л. Л. Афанасьева. Изд. 3, переработ. и доп. М., Машгиз, 1959, 892 стр. с илл.; 1 л. черт. Авт.: Л. Л. Афанасьев, В. А. Илларионов, Н. Э. Струве и К. С. Шестопалов.

На многотомное издание, т. е. издание, которое состоит из двух или более томов и представляет собой по содержанию единое целое (издание ограничено определенным числом томов), составляется одно сводное описание. Так как при оформлении выходных сведений многотомников определенные данные повторяются

в каждом томе, сводное описание делится на две части. В общей части сводного описания один раз указываются все сведения, повторяющиеся в каждом томе, а в спецификации — лишь особенности, отличающие один том от другого.

Образец сводного описания многотомника

Общая часть

Теория авиационных двигателей. Учебник для авиац.-техн. учебн. заведений. Ч. 1—2. М, Воениздат, 1960.
Библиогр. в конце каждой части.

Спецификация

Ч. 1. Мадорский Я. Ю. и Ровинский Э. В. Основы термодинамики и газовой динамики. 211 стр. с илл.
Ч. 2. Востриков С. И., Зуев Л.Н., Кузнецов В. И. и др. Теория реактивных двигателей. Под ред. И. В. Котляра. 282 стр. с илл.

Метод сводного описания применяется также к серийным и периодическим изданиям.

Официальные материалы директивного и отчетного характера, ведомственные нормативные документы, производственные и технологические инструкции, информационно-справочные и другие материалы учреждений и организаций описываются под коллективным автором.

Образец описания под коллективным автором

Московский энергетический институт.
Сборник работ аспирантов за 1964 г. М., 1965.
125 стр. с черт.

Аналитическое описание применяется для отдельного отражения частей изданий или документов. Часть какого-либо издания, например статья в сборнике или журнале, сама по себе является научным документом. Поэтому статья описывается, как и любой документ, под автором и заглавием. Но вместо выходных данных в этом случае указываются сведения об издании, в котором данная статья помещена. Эти сведения составляются по-разному, в зависимости от того, в каком издании помещена статья — в сборнике, журнале, продолжающемся издании или газете.

Образец аналитического описания части издания

Покровский Г.

Длительный полет человека по орбите. — «Вестник воздушн. флота», 1961, № 9, стр. 80—84.

Совсем иную форму имеет аналитическое описание части документа, например его главы или раздела. Глава или раздел документа не может восприниматься самостоятельно, в отрыве от этого документа. Поэтому описание части документа производится на общих основаниях под автором и заглавием документа. Но вместо количественной характеристики всего издания в целом в аналитическом описании указывается номер и заглавие соответствующей главы или раздела, а также страницы, которые он занимает в документе.

Образец аналитического описания части документа

Новикова Е. А.

Описание произведений печати и организация алфавитного каталога в технических библиотеках. Метод. пособие. М., 1964.

Гл. 7. Особенности построения заголовка описания изданий разных учреждений и организаций. Стр. 67—95.

Некоторыми особенностями отличаются правила описания специальных видов технической литературы. В нашей стране стандарты описываются по правилам, соответствующим международной практике и применяемым для печатных карточек, издаваемых Государственным комитетом стандартов, мер и измерительных приборов при Совете Министров СССР. Слева вверху помещается индекс УДК, а под ним — шифр группы по классификатору «Указателя государственных стандартов»; в правом верхнем углу приводится географический определитель по УДК и наименование страны, в которой был принят данный стандарт. Ниже помещаются заглавие и подзаголовочные данные стандарта. С новой строки приводится аббревиатура, обозначающая вид стандарта, его национальную принадлежность или организацию, разработавшую и принявшую данный стандарт (для иностранных стандартов). Например, ГОСТ — государственный общесоюзный стандарт, BS — British standard, DIN — Deutsche Industrie Norme, ASME — American Society of Mecanical Engineers. После этой аббревиатуры указывается номер стандарта,

две последние цифры года его утверждения, сведения о ранее действовавшем стандарте, если данный стандарт вводится взамен другого стандарта, а также дата введения данного стандарта. В правом нижнем углу указывается сокращенное наименование учреждения, утвердившего стандарт.

Образец описания стандарта

621.432 (47) СССР
Группа Д29

Поршневые двигатели внутреннего сгорания. Воздухоочистители. Методы стендовых безмоторных испытаний.
ГОСТ 8002-62, взамен ГОСТ 8002-56. Введ. 1/I—1964 г.

ГКСМ и ИП

Описания изобретений к авторским свидетельствам и патентам каталогизируются под именем изобретателя или патентообладателя (отдельного лица или учреждения). В библиографическом описании указываются также: название изобретения, номер авторского свидетельства или патента, сведения о заявке, даты опубликования и выдачи авторского свидетельства или патента, его классификационный индекс и другие сведения.

Образец описания авторского свидетельства

| | |
|---|---|
| 42 т. 14 | |
| 145065 | Бруснецов Н. П. и Маслов С. П.<br>Троичный счетчик. Авт. свид. 145065.<br>Заявл. 7/II—1961 г. в Комитет по делам изобретений и открытий при Совете Министров СССР за № 696640/26.<br>Опубл. в «Бюлл. изобретений», 1962, № 4. |

Описание технических каталогов производится по-разному в зависимости от характера изданий и типа опубликовавшего их учреждения или предприятия. Технические каталоги министерств и главных управлений, отражающие однотипную продукцию различных предприятий, описываются под заглавием. Каталоги трестов и промышленных предприятий, отражающие разнообразную продукцию одного треста или завода, описываются под коллективным автором. Каталоги-справочники описываются по правилам описания книг, чаще всего — под фамилией автора или составителя.

Образец описания технического каталога под заглавием

| Каталог моделей обуви. Киев, 1961. |
|---|
| 312 стр. с илл. (Киевск. дом моделей). |

Образец описания технического каталога под коллективом

| Горьковский автомобильный завод |
|---|
| Каталог запасных частей грузовых автомобилей ГАЗ-51А, ГАЗ-51Ж, ГАЗ-51П, ГАЗ-63, ГАЗ-63А, ГАЗ-63Д, ГАЗ-93 и ГАЗ-93А. М., Машгиз, 1960. |
| 339 стр. с илл. |

Образец описания технического каталога-справочника

| Крутов М. И. и Самоданова В. М. |
|---|
| Каталог запасных частей к машинам по возделыванию льна и конопли. М., Сельхозгиз, 1959. |
| 148 стр. с илл. |

### *Библиографическое описание в научно-информационной деятельности*

Правила библиографического описания до сих пор составлялись в расчете лишь на традиционные формы каталогов, картотек и указателей, без учета возможностей машинного поиска информации. Специфические требования научно-информационной деятельности и применение средств механизации и автоматизации оказывают определенное влияние на теорию и практику описания. Однако нет оснований предполагать, что основные принципы книгоописания, сложившиеся в течение многовековой практики, должны подвергнуться коренной ломке.

Библиографическое описание в научно-информационной деятельности выполняет две основные функции — адресную и сигнальную. Функция адресности выполняется благодаря тому, что описание содержит необходимые и достаточные сведения для отыскания и отождествления определенных научных документов по их основным библиографическим признакам: фамилии автора и других лиц, принимавших участие в создании документа, наименованию учреждения, ответственного за содержание и издание документа, первым словам заглавия документа. В этом отно-

шении состав библиографического описания удовлетворяет современным требованиям. Любой документ, независимо от того, издан он отдельно или является частью какого-либо издания, может быть отыскан и опознан по его описанию. Применение машинной техники позволит увеличить число признаков, по которым ведется поиск документов.

При использовании традиционных библиотечно-библиографических средств информационного поиска — каталогов, картотек и указателей — обычно приходится ограничивать число авторов, имена которых служат характеристиками для поиска документов. Необходимость в таком ограничении отпадет, когда информационный поиск будет вестись при помощи соответствующих машин. Это даст возможность отражать в описании не только имена всех авторов, но и имена других лиц, участвовавших в создании документа — составителей, авторов отдельных глав или статей, редакторов, переводчиков и т. д.

Сигнальная функция описания заключается в том, что оно может использоваться в качестве средства оповещения о появлении документа, имеющего определенное научное содержание. Из всех элементов описания для этой цели могут быть использованы лишь заглавия и подзаголовочные данные, а также отчасти заглавия серий в надзаголовочных данных. Эти данные не всегда в достаточной мере раскрывают содержание документа. Кроме того, заглавия, которые даются документам их авторами, составителями или издателями, иногда не соответствуют содержанию документов, а необходимость в сохранении точного адреса документа не позволяет произвольно менять его заглавие. Тем не менее в реферативных журналах все чаще практикуется замена заглавий, не соответствующих содержанию статей, другими, более точными названиями. Необходимо отметить, что такая практика применима лишь в отношении журнальных статей, заглавия которых не несут большой адресной нагрузки.

Особого внимания заслуживает вопрос о последовательности элементов описания. Традиционный их порядок — автор, заглавие, подзаголовочные и выходные данные, количественная характеристика, надзаголовочные данные и примечания — принят во всех каталогах и библиографических указателях и в большинстве реферативных изданий. В некоторых реферативных журналах, в частности в «Реферативном журнале ВИНИТИ», в «Chemical Abstracts», «Biological Abstracts» и других, заглавие выносится всегда на первое место. Установление такой последовательности элементов описания объясняется тем, что читателю, просматривающему реферативный журнал, важно знать в первую очередь содержание документов. По-видимому, внедрение машинных методов в составление каталогов и библиографических указателей будет влиять на внешний вид описания, состав и последовательность его элементов. Однако не следует забывать, что основные

принципы описания разрабатывались с учетом практических потребностей научных работников и рассчитаны на комплексное использование всех видов научных документов. Поэтому произвольное изменение методов описания документов может повести к затруднениям в научно-информационной деятельности.

Решение данной проблемы лежит за пределами области библиографического описания и больше касается издательского процесса. Почти во всех странах существуют стандарты на оформление библиографических сведений в издательской продукции, но эти стандарты не всегда соблюдаются. Многие документы не содержат сведений, необходимых для правильного и полного их описания. Особенно тщательного внимания заслуживают заглавия. Заглавие должно носить описательный характер, быть кратким и вместе с тем точно раскрывать содержание документа. Если составить такое заглавие не удается, то оно должно сопровождаться соответствующим подзаголовком. Ввиду все более широкого применения пермутационных указателей желательно, чтобы в заглавие входили ключевые слова, отражающие смысловое содержание документа.

## КЛАССИФИКАЦИЯ (СИСТЕМАТИЗАЦИЯ)

Классификация (систематизация) документальной информации является одним из важнейших видов аналитико-синтетической переработки и заключается в группировке научных документов по отраслям знаний, отдельным проблемам, предметам или по каким-либо формальным признакам (например, расстановка документов или изданий в алфавитном порядке имен их авторов или их заглавий). В более узком смысле библиотечно-библиографическая классификация — это распределение документов по отраслям знаний в соответствии с содержанием этих документов. Такие классификации документов в той или иной мере связаны с классификациями наук, так как их основные классы обычно соответствуют определенным отраслям знаний, а характер дальнейшего деления этих классов соответствует структуре самих отраслей знания.

Вместе с тем библиотечно-библиографическая классификация не идентична классификации наук. Основные отличия классификации наук от библиотечно-библиографической классификации определяются главным образом сугубо практическим характером последней, а также особенностями классифицируемых объектов — книг, журнальных статей и других научных документов. Помимо делений по содержанию документов, библиотечно-библиографическая классификация содержит также деления по видам изданий (книги, периодические издания, статьи, специальные виды технических изданий), по их назначению (научно-популярная, учебная, специальная литература), по языку их текста (на рус-

ском языке, на языках других народов СССР, на иностранных языках) и т. п.

Схемы библиотечно-библиографической классификации обычно издаются в виде таблиц, которые состоят из двух частей: основных таблиц и таблиц определителей или типовых делений. В основных таблицах все отрасли знания и их разделы расположены в логической последовательности, причем деление каждый раз производится только по одному основанию. Каждая рубрика схемы содержит ряд подчиненных ей рубрик, образуя определенную иерархию. Из-за этих особенностей традиционные библиотечно-библиографические схемы иногда называют «линейно-иерархическими». В таблицах определителей или перечнях типовых делений отражаются общие повторяющиеся для многих предметов признаки (например, признак места и времени, к которым относится содержание документа, вида издания и др.)

Каждому разделу библиотечно-библиографической классификации присваивается условное обозначение, называемое индексом. Для этого используются индексы цифровые, буквенные и смешанные. Достоинства цифровых индексов состоят в том, что порядковая последовательность цифр более очевидна, чем алфавитная последовательность букв, что любые сочетания цифр легко произносятся как определенные номера или числа и что арабские цифры понятны всем народам, независимо от языка, на котором они говорят, и письменности, которой они пользуются.

Вместе с тем цифровые индексы обладают существенным недостатком — ограниченностью базы. Так как цифр всего только десять (от 0 до 9), то для обозначения сложных или очень узких понятий приходится использовать громоздкие индексы, состоящие из большого числа знаков. Для устранения этого недостатка часто прибегают к смешанным индексам, в которых, наряду с цифрами, применяются также буквы того или иного алфавита.

По структуре индексы могут быть номерными и ступенчатыми. В качестве номерных индексов используются порядковые номера подразделений классификации в общем числовом ряду. Такие цифровые индексы отличаются низкой мнемоничностью, не отражают иерархии разделов классификации, а также затрудняют дальнейшее деление этих разделов. Поэтому в большинстве современных библиотечно-библиографических классификаций применяются ступенчатые индексы, которые могут быть и цифровыми, и буквенными, и смешанными. Ступенчатые индексы отражают логическую структуру классификации, так как каждый ее раздел обозначается каким-либо одним символом, а все деления этого раздела — двумя символами, из которых первый обозначает раздел, а второй — соответствующий подраздел этого раздела. Применение ступенчатых индексов позволяет производить неограниченную детализацию классификационной схемы.

## *Из истории библиотечно-библиографических классификаций*

История библиотечно-библиографических классификаций восходит к глубокой древности, и это естественно, так как потребность в такой классификации возникла одновременно с появлением письменных документов. До нас дошло очень немного сведений о классификациях, применявшихся в древних библиотеках. По всей вероятности, в античные времена такие классификации были разработаны достаточно детально. Об этом можно судить по обширности библиотек и высокому уровню классификации наук того времени, а также по некоторым косвенным свидетельствам, содержащимся в литературных памятниках античности.

На библиотечные классификации раннего средневековья известное влияние оказала распространенная в то время система «семи свободных искусств». Эта система состояла из двух комплексов наук, изучавшихся в средневековой школе: «тривиума» (грамматики, диалектики и риторики) и «квадривиума» (арифметики, геометрии, музыки и астрономии). К концу XV в. в университетских библиотеках начали применять группировку книг по содержанию в соответствии с существовавшими тогда в большинстве университетов четырьмя факультетами: философским, медицинским, юридическим и богословским. Это послужило толчком к возникновению так называемых «факультетских систем» классификации, пользовавшихся широкой популярностью на протяжении ряда столетий, вплоть до XIX в. Факультетские системы оказали влияние на такие выдающиеся для своего времени классификации, как классификации швейцарского ученого и библиографа К. Геснера (1548 г.) и немецкого философа-идеалиста Г. Лейбница (1700 г.)

Дальнейшее развитие библиотечно-библиографической классификации проходило под воздействием идей английского философа Френсиса Бэкона (1561—1626 гг.), которого К. Маркс назвал «истинным родоначальником *английского материализма* и вообще *опытных* наук *новейшего времени*» [7]. Созданная им в начале XV II в. классификация наук группировала знания в соответствии с идеалистической традицией по «способностям человеческого духа». «Память» определяла возникновение истории, «Воображение» — поэзии, а «Разум» — философии или собственно науки. При всей условности такого деления классификация Ф. Бэкона включала новые отрасли знания и представляла собой значительное событие для науки того времени.

В XIX в. библиотечно-библиографические классификации стали широко разрабатываться в России. Большой интерес представляет схема ученого-натуралиста П. Г. Демидова, составлен-

---

[7] К. Маркс и Ф. Энгельс. Сочинения, т. 3, стр. 157.

ная им для каталога личной библиотеки. В этой схеме наиболее детализированными были разделы естественных наук и технологии. Заслуживает внимания схема, опубликованная в 1809 г. А. Н. Олениным и предназначенная для петербургской Публичной библиотеки, а также оригинальная схема Ф. Ф. Рейсса, предложенная им для библиотеки Московского университета (1926 г.). Хотя в основу этой схемы и был положен формальный принцип дихотомии, т. е. последовательного деления объема каждого понятия на два противоречащих друг другу видовых понятия, Рейссу удалось образовать в схеме классификации большое число делений для различных типов изданий и разработать удобную систему ступенчатых индексов.

Самостоятельную схему классификации, в основу которой было положено условное деление наук «по потребностям человека», создал для библиотеки Казанского университета профессор этого университета К. К. Фойгт (1843 г.). Одной из лучших в первой половине XIX в. заслуженно считается классификация выдающегося русского натуралиста академика К. Э. Бэра, которую он разработал для иностранного отделения библиотеки Академии наук в Петербурге (1841 г.). Эта классификация была одной из первых попыток расположить науки в последовательности, отражающей историю развития мира: науки о неорганической природе, науки об органической природе и науки о человеке и человеческом обществе.

В библиотеках Западной Европы в этот период особенно широко применялись схемы классификации Ж. Ш. Брюне (1810 г.) и А. Э. Шлейермахера (1847г.). Первая из них представляла собой одну из поздних разновидностей так называемой «французской системы», которая в течение почти двух столетий использовалась в библиографии и книготорговле Франции. Вторая, особенно популярная в библиотеках немецких университетов, была создана на основе старой «факультетской системы».

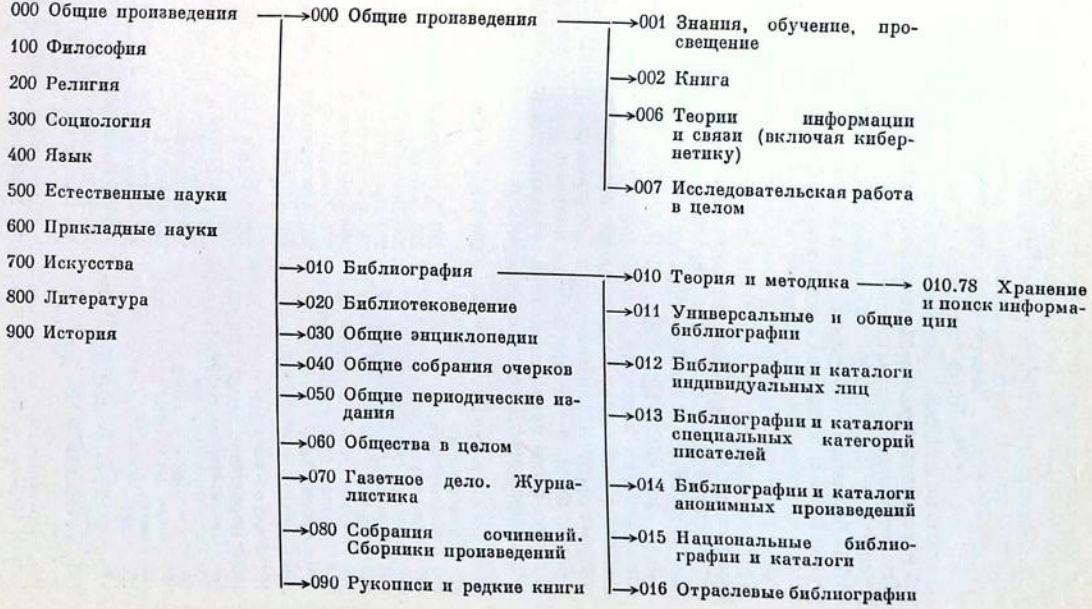
### «Десятичная классификация» М. Дьюи (1876 г.)

Дальнейшее развитие библиотечно-библиографическая классификация получила в связи с успехом библиотечного дела в США во второй половине XIX в. Мельвиль Дьюи (1851—1931 гг.), впоследствии известный американский библиотековед, в 1876 г. опубликовал схему классификации книг, предназначенную для библиотеки небольшого колледжа. В ней он впервые применил десятичный принцип деления, который в дальнейшем получил широкое распространение в классификационных схемах.

В отношении классификации наук схема М. Дьюи не была оригинальной. Распространено мнение, что он использовал в обратном порядке традиционную последовательность расположения

Таблица 4

**Схема деления отдела «Общие произведения» в «Десятичной классификации» Дьюи (изд. 16)**



отделов, восходившую к классификации наук Ф. Бэкона, что первые шесть отделов его схемы соответствуют третьей группе у Ф. Бэкона («Философия» или «Науки»), седьмой и восьмой отделы — второй группе («Поэзия»), девятый отдел — первой группе («История»). Следует отметить, что это мнение не совсем точно, так как М. Дьюи был далек от классификационных идей английского философа. Последовательность классов в «Десятичной классификации» он заимствовал из широко распространенных предшествующих классификаций, в частности из классификации Брюне. Для Дьюи логический принцип расположения отраслей знания был менее существенным, чем формальный принцип последовательного деления всей совокупности знаний и каждой отрасли знания на десять частей. Поэтому в расположении отделов его схемы нарушена естественная связь между науками, схема не отражает реальной картины мира не только в современном понимании, но и в трактовке того времени.

Своим успехом «Десятичная классификация» Дьюи обязана главным образом ступенчатым индексам, в которых используются только арабские цифры (табл. 4). Каждый основной раздел этой классификации обозначен одной цифрой; при дальнейшем его делении к этой цифре присоединяется сначала вторая, а затем и третья цифра. К однозначным и двузначным числам М. Дьюи добавил нули, с тем чтобы везде получать только трехзначные индексы. Ниже приводится схема основных делений «Десятичной классификации» Дьюи с их цифровыми индексами [8]:

000 Общие произведения
010 Библиография
020 Библиотековедение
030 Общие энциклопедии
040 Собрания сочинений
050 Общие периодические издания
060 Общества
070 Газетное дело. Журналистика
080 Собрание сочинений. Сборники произведений
090 Рукописи и редкие книги

100 Философия
110 Метафизика
120 Метафизические теории
130 Разделы психологии
140 Философские системы
150 Общая психология
160 Логика
170 Этика
180 Древняя и средневековая философия
190 Современная философия

200 Религия
210 Естественное богословие
220 Библии
230 Доктрины богословия
240 Практическое богословие
250 Пастырское богословие
260 Христианская церковь
270 История христианской церкви
280 Христианские церкви и секты
290 Другие религии

300 Социальные науки
310 Статистика
320 Политические науки
330 Экономика
340 Право
350 Общественное управление
360 Общественное благосостояние
370 Образование
380 Торговля и связь
390 Обычаи и фольклор

[8] Dewey Decimal classification and relative index. Dewised by Melvil Dewey. 16 th ed. v. 1—2. New York, Forest Press, 1958.

400 Язык
410 Сравнительная лингвистика
420 Английский и англо-саксонские языки
430 Германские языки
440 Французский, провансальский, каталанский языки
450 Итальянский, румынский языки
460 Испанский, португальский языки
470 Латинский и другие италийские языки
480 Древнегреческий и новогреческий языки
490 Другие языки

500 Естественные науки
510 Математика
520 Астрономия
530 Физика
540 Химия и смежные науки
550 Науки о Земле
560 Палеонтология
570 Антропология и биология
580 Ботанические науки
590 Зоологические науки

600 Технические и прикладные науки
610 Медицинские науки
620 Техника
630 Сельское хозяйство
640 Домоводство
650 Делопроизводство. Управление хозяйством
660 Химическая технология
670 Отрасли производства
680 Другие отрасли производства
690 Строительное дело
700 Искусства
710 Садоводство
720 Архитектура
730 Скульптура
740 Рисунок и декоративное искусство
750 Живопись
760 Печатная графика
770 Фотография
780 Музыка
790 Развлечения

800 Литература
810 Американская литература на английском языке
820 Английская и древнеанглийская литературы
830 Литература на немецком языке
840 Французская, провансальская, каталанская литературы
850 Итальянская, румынская литературы
860 Испанская, португальская литературы
870 Латинская и другие италийские литературы
880 Древнегреческая и новогреческая литературы
890 Другие литературы

900 История
910 География, путешествия
920 Биографии
930 Древняя история
940 Европа
950 Азия
960 Африка
970 Северная Америка
980 Южная Америка
990 Другие части света

Особое значение М. Дьюи придавал тому, чтобы его схема была удобна для практического использования. Для обозначения формы документов и видов изданий М. Дьюи разработал фиксированный список делений, индексы которых могут непосредственно присоединяться к индексам любого отдела классификации. Этот список имеет следующий вид:

01 Теория, полезность и т. д.
02 Компендиумы (краткие руководства)
03 Словари
04 Сборники
05 Периодические издания
06 Общества
07 Образование, изучение
08 Смешанные произведения
09 История

Таким образом, в подотделах 531 — «Механика» и 622 — «Горное дело» могут быть образованы следующие индексы»:

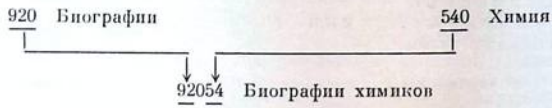
| | |
|---|---|
| 53101 Теория механики | 62201 Теория горного дела |
| 53102 Руководства по механике | 62202 Руководства по горному делу |
| 53103 Словари по механике | 62203 Словари по горному делу |
| 53109 История механики | 62209 История горного дела |

Для характеристики документов по языковому или географическому признакам М. Дьюи предложил использовать окончания соответствующих индексов отделов 400 — «Филология» и 900 — «История». В подразделах 390 — «Фольклор» и 550 — «Геология» по этому методу можно образовать следующие индексы:

| | |
|---|---|
| 39020 Английский фольклор | 55040 Геология Европы |
| 39030 Немецкий фольклор | 55050 Геология Азии |
| 39040 Французский фольклор | 55060 Геология Африки |

Для отнесения библиографических документов к той или иной отрасли знания М. Дьюи предложил использовать в подотделе 920 — «Библиографии» первые цифры индексов соответствующих подразделений своей классификации.

Например:




Примененные в «Десятичной классификации» Дьюи методы были первым шагом на пути к созданию типовых делений или определителей, играющих большую роль в современных схемах библиотечно-библиографической классификации. Не менее важным нововведением было и составление к таблицам классификации алфавитно-предметного указателя, до сих пор применявшегося лишь в энциклопедиях и библиографиях с систематическим расположением материала. М. Дьюи считал, что подобный указатель дает возможность найти в классификации любое подразделение, даже если занимаемое им в схеме место не имеет достаточно логического обоснования.

«Десятичная классификация» Дьюи широко используется в библиотеках США и Великобритании. Можно считать, что она применяется в 90% всех библиотек этих стран. В 1958 г. в США вышло 16-е издание этой классификации, в 1965 г. — 17-е издание. В последних ее изданиях особенно явно проявилась реакционность идеологии составителей этой классификации, их стремление затушевать острые проблемы международной и внутренней политики, а также нежелание противодействовать распространению антинаучной литературы. Известные пороки в научном и

идеологическом отношении содержались уже в первом издании «Десятичной классификации» Дьюи: религия в ней помещена почти в самом начале схемы (отдел 200), история оторвана от общественных наук (300) и философии (100), литература (800) — от филологии (400). В прикладных науках (600) объединены такие разные отрасли, как медицина, сельское хозяйство и техника.

Ограниченность цифровой базы индексов придает еще бо́льшую искусственность «Десятичной классификации» Дьюи, так как в действительности различные отрасли науки могут делиться не на десять, а на значительно большее число подразделений. Однако практические удобства десятичного принципа при его творческом, неформальном использовании настолько велики, что он и в настоящее время не утратил своего значения. Достаточно сказать, что большинство применяемых сейчас схем классификации, включая и советские схемы, в той или иной мере используют этот принцип.

## «Растяжимая классификация» Ч. Кеттера (1879 г.)

Другой видный теоретик американского библиотековедения Чарльз Кеттер (1837—1903 гг.) при организации фондов библиотеки Бостонского Атенеума [9] создал свою схему, в которой попытался устранить недостатки десятичной классификации. В отличие от М. Дьюи Ч. Кеттер придавал большое значение тому, чтобы библиотечно-библиографическая классификация соответствовала классификации наук и чтобы в ней соблюдалась логическая последовательность расположения разделов. В качестве основания для деления всей совокупности знаний Ч. Кеттер выбрал эволюционный принцип развития природы. Несмотря на то, что этот принцип понимался им упрощенно и с идеалистических позиций, сам факт связи библиотечной классификации с эволюционным учением имел весьма положительное значение. В его схеме устранена логическая непоследовательность «Десятичной классификации» Дьюи: сближены отделы истории и общественных наук, языкознания и литературы; медицина выделена из прикладных наук и расположена непосредственно за естественными науками; строительное дело и промышленное производство разделены.

Основной особенностью «Растяжимой классификации» Кеттера было создание нескольких вариантов таблиц, рассчитанных на библиотеки с фондами разной величины (от ста до десяти миллионов томов). Каждый из вариантов содержал одни и те же

[9] Бостонский Атенеум — одно из наиболее известных культурно-просветительных учреждений США в XIX в.

основные классы, но отличался от предыдущего большей дробностью деления этих классов. Ч. Кеттер опубликовал шесть вариантов таблиц, седьмой остался незаконченным. Идея расширимости (растяжимости) была заложена и в «Десятичной классификации» Дьюи, которая допускает бесконечное деление классов без изменения общей структуры схемы. Однако только у Ч. Кеттера эта идея получила практическое воплощение.

Стремясь преодолеть ограниченность цифровой базы десятичной классификации, Ч. Кеттер применил ступенчатые буквенные индексы, составленные из прописных и строчных букв латинского алфавита. Основные деления его схемы видны из следующей таблицы [10].

A Общие произведения, охватывающие несколько классов
B Философия
Br Религия
C Христианство и иудаизм
D Церковная история
E Биографии
F История
G География и путешествия
H Социальные науки
Hc Экономика. Политическая экономия
I Демотика. Социология
Ip Образование
J Государственное управление
Ja Гражданские науки. Политические науки
K Законодательство и пр.
L Науки и искусства вместе
La Науки (естественные)
Lb Математика
Lh Физика. Естественная философия
Lo Химия
Lp Астрономия
M Естественная история
Mb Микроскопия
Mc Геология
Md Минералогия. Литология
Me Кристаллография
Mg Физиография. Физическая география
Mp Палеонтология
Mu Биология
N Ботаника
O Зоология
Pw Антропология и этнология
Py Этнография (расы людей)
Q Медицина
R Полезные искусства. Техника
Rd Горное дело
Re Металлургия
Rf Сельское хозяйство
Rg Ветеринария
Rj Химическая технология
Rt Электротехника
Ry Домашнее хозяйство
Rz Пища и кулинария
S Конструктивные искусства (техника и строительство)
Sg Строительное дело
Sj Санитарная техника
Sl Гидравлическая техника
St Транспорт и связь
Su Дороги, шоссе, улицы
Sv Железные дороги
T Мануфактуры и ремесла
U Военное искусство
Un Морское дело
Uu Суда, судостроение и пр.
V Развлечения. Спорт. Игры. Празднества
Vs Гимнастика. Физическое воспитание
Vt Театр
Vv Музыка
W Искусство
We Ландшафтное садоводство
Wf Архитектура
Wj Скульптура
Wk Гравирование и вытачивание
Wl Рисование. Графические искусства
Wp Живопись
Wq Гравюра
Wr Фотография
Ws Декоративное искусство.

[10] Е. И. Шамурин. Очерки по истории библиотечно-библиографической классификации, т. 2. М., 1959, стр. 193—194.

Орнамент
Wt Мозаика
Wu Рукоделия и текстильные украшения
Wv Костюмы и их принадлежности
Ww Мебель
Wx Ювелирное дело
Wy Работа по металлу
X Филология
Y Английская и американская литературы
Yi Юношеская литература
Z Искусство книги
Za Авторство. Риторика
Zd Письмо
Zh Печатание
Zk Переплетное дело
Zi Издательство и книготорговля
Zn Частные библиотеки
Zp Публичные библиотеки
Zu Библиография
Zw Предметная библиография
Zx Выбор и методы чтения
Zy Литературная история
Zz Национальная библиография

Дальнейшее развитие в «Растяжимой классификации» Кеттера [11] получила идея М. Дьюи о создании стандартных индексов для обозначения видов документов и географических пунктов. Ч. Кеттер впервые детально разработал и ввел в классификацию таблицы определителей формы и географических определителей. Эти определители обозначались арабскими цифрами и присоединялись к индексу или с помощью точки (определители формы) или непосредственно (географические), например:

Важной заслугой Ч. Кеттера является и то, что он позаботился об обеспечении алфавитного порядка расположения документов внутри каждого деления схемы классификации. Для облегчения расстановки документов в алфавитном порядке Ч. Кеттер составил так называемые «Авторские таблицы». В этих таблицах все фамилии, начинающиеся с одних и тех же слогов, получили одинаковое обозначение, состоящее из начальной буквы фамилии и порядкового номера начального слога в алфавитном списке слогов, начинающихся с соответствующей буквы. Эти таблицы, получившие название «кеттеровских», применяются сейчас во всех библиотеках с систематической расстановкой фондов.

Фрагмент трехзначных авторских таблиц

| | | | | | |
|---|---|---|---|---|---|
| Пане | 161 | Раев | Пант | 166 | Раец |
| Пани | 162 | Раевс | Панф | 167 | Раеч |
| Панк | 163 | Раевски | Панч | 168 | Раеш |
| Панкр | 164 | Раек | Пао | 169 | Раж |
| Пано | 165 | Раен | | | |

[11] C. A. Cutter. Expansive classification. P. 1—2. Boston, 1891—1898.

Таблица 5

**Схема деления отделов «Общие произведения. Полиграфия» и «История» в «Классификации Библиотеки конгресса США»**

A Общие произведения. Полиграфия
- → AC Коллекции. Серии. Сборники.
- → AE Энциклопедии
- → AG Общие справочники (кроме энциклопедий)
- → AI Указатели
- → AM Музеи → 200—275 Коллекции, коллекционирование и частные кабинеты
- → AN Газеты
- → AP Журналы
- → AS Общества. Академии
- → AY Ежегодники. Альманахи
- → AZ Общая история знаний и образования

B Философия. Религия

C История. Вспомогательные дисциплины
- → CB История цивилизации

D История и топография (кроме Америки)
- → CC Древности (в целом)
- → CD Архивы. Дипломатика → 5001—6471 Сфрагистика

E Америка и США

F США (отдельные местности) и Америка (кроме США)
- → CE Хронология
- → CT Нумизматика

G География. Антропология
- → CN Эпиграфика. Надписи
- → CR Геральдика

H Общественные науки
- → CS Генеалогия

J Политические науки
- → CT Биография

K Право

L Воспитание

M Музыка

N Изящные искусства

P Язык и литература

Q Науки

R Медицина

S Сельское хозяйство

T Техника

U Военные науки

V Морские науки

Z Библиография и библиотековедение

Несмотря на большие теоретические и практические достоинства «Растяжимой классификации» Кеттера, она не получила широкого распространения и применяется лишь в нескольких десятках крупных публичных и университетских библиотек США. Это объясняется тем, что ко времени опубликования более детализированных вариантов «Растяжимой классификации» Кеттера (90-е годы XIX в.) большинство американских библиотек уже применяло «Десятичную классификацию» Дьюи.

## «Классификация Библиотеки конгресса США» (1901 г.)

Начиная с 1901 г. в течение 40 лет отдельными томами публиковалась подробная классификация, разрабатываемая в одной из крупнейших мировых библиотек — в Библиотеке конгресса США. В целом она занимает 28 томов общим объемом около 6 тыс. страниц. Эта классификация отличается большой детализацией, она применяется в сотнях крупных научных и специальных библиотек США, ее индексы приводятся на печатных карточках Библиотеки конгресса (наряду с индексами «Десятичной классификации» Дьюи). Классификация разрабатывалась группой специалистов под руководством американского библиотековеда Ч. Мартела.

По существу «Классификация Библиотеки конгресса» представляет сумму мало связанных между собой классификаций отдельных отраслей знания. Поскольку работа над каждым разделом классификации велась в отрыве от разработки других ее разделов, в ней нет единства методических решений и технических приемов. Сильное влияние на детализацию отдельных разделов этой классификации оказал весьма неоднородный состав фондов Библиотеки конгресса. Для индексации таблиц применены смешанные буквенно-цифровые ступенчато-номерные индексы (табл. 5). Основные разделы классификации обозначены прописными буквами латинского алфавита, а дальнейшие их подразделения — трехзначными номерами, каждый из которых указывает лишь место подразделения в разделе, но не отражает его логической подчиненности.

Схема «Классификации Библиотеки конгресса США» включает следующие основные разделы [12]:

A Общие произведения. Полиграфии
AZ Общая история знаний и образования
B Философия. Религия
C История. Вспомогательные исторические дисциплины
CT Биографии (общие)
D История и топография (кроме Америки)

[12] Library of Congress. Classification. Outline scheme of classes. Washington, Government printing office, 1922.

E Америка и США
E 11—143 Америка (в целом)
E 31—50 Сев. Америка (в целом)
E 51—99 Индейцы Сев. Америки
E 101—135 Открытие Америки
E 151—770 США
E 151—185 Общая история
E 185 Негры в США
E 186—200 Колониальный период
E 201—298 Революция
E 351—364 Война 1812 г.
E 401—415 Война с Мексикой
E 441—453 Рабство
E 458—655 Гражданская война
E 714—735 Война с Испанией
F США (отдельные местности) и Америка (кроме США)
G География. Антропология
GA География, путешествия
GF Антропогеография
GN Антропология. Соматология. Этнология. Этнография
GR Фольклор. Культура и цивилизация. Обычаи и нравы
GV Спорт и развлечения. Игры
H Общественные науки
HA Статистика

Экономика

HB Экономическая теория
HC Экономическая история (Национальное производство)
HE Транспорт и связь
HF Торговля
HG Финансы

Социология

HM Социология, общая
HX Социализм. Коммунизм. Анархизм
J Политические науки
JC Теория государства

Конституционная история и управление

JE Общее
JK США
JS Местное управление
JV Колонии и колонизация. Эмиграция и иммиграция
JX Международное право
K Право
L Воспитание
M Музыка
N Изящные искусства
NA Архитектура
NB Скульптура и родственные искусства
NC Графические искусства вообще
ND Живопись
NE Гравюры
NF Фотография (как искусство)
NK Прикладные искусства
P Язык и литература
P Филология. Лингвистика
PN—PV История литературы. Литература
PZ Художественная литература
Q Науки
QA Математика
QB Астрономия
QC Физика
QD Химия
QE Геология
QH Естественная история
QK Ботаника
QL Зоология
QM Анатомия человека
QP Физиология
QR Бактериология
R Медицина
S Сельское хозяйство. Растениеводство и животноводство
SF Животноводство. Ветеринария
SH Рыбоводство и рыболовство
SK Охота. Пушной промысел. Охрана животных
T Техника
T Техника в целом

Техническая и строительная группа

TA Гражданская техника
TC Гидравлическая техника
TD Санитарная и муниципальная инженерия
TE Дороги и мостовые
TF Железные дороги
TG Мосты и покрытия
TH Постройка зданий

Механическая группа

TJ Механическая техника
TK Электротехника и электропромышленность
TL Средства передвижения. Аэронавтика

Химическая группа

TN Горное дело
TP Химическая технология
TR Фотография

Смешанная группа
TS Мануфактуры
TT Ремесла
TX Домоводство
U Военные науки
V Морские науки
Z Библиография и библиотековедение
Z4 История книги и книгопроизводства
Z 40—115 Письменность
Z 116—550 Книгопроизводство и книготорговля
Z 551—657 Авторское право. Копирайт
Z 665—997 Библиотеки и библиотековедение
Z 999—1000 Цены на книги
Z 1001—9000 Библиография

## «Универсальная десятичная классификация» (1905—1907 гг.)

Наиболее значительной из всех библиотечно-библиографических классификаций, созданных на рубеже прошлого и настоящего столетий, шире всего распространенной и больше всего применяемой во всех странах мира, следует признать «Универсальную десятичную классификацию». Ее создание связано с именами выдающихся бельгийских библиографов Поля Отле (1868—1944) и Анри Лафонтена (1854—1943). Созванная по их инициативе в 1895 г. международная библиографическая конференция учредила две организации: Международное бюро библиографии и Международный библиографический институт, которые впоследствии объединились под вторым названием.

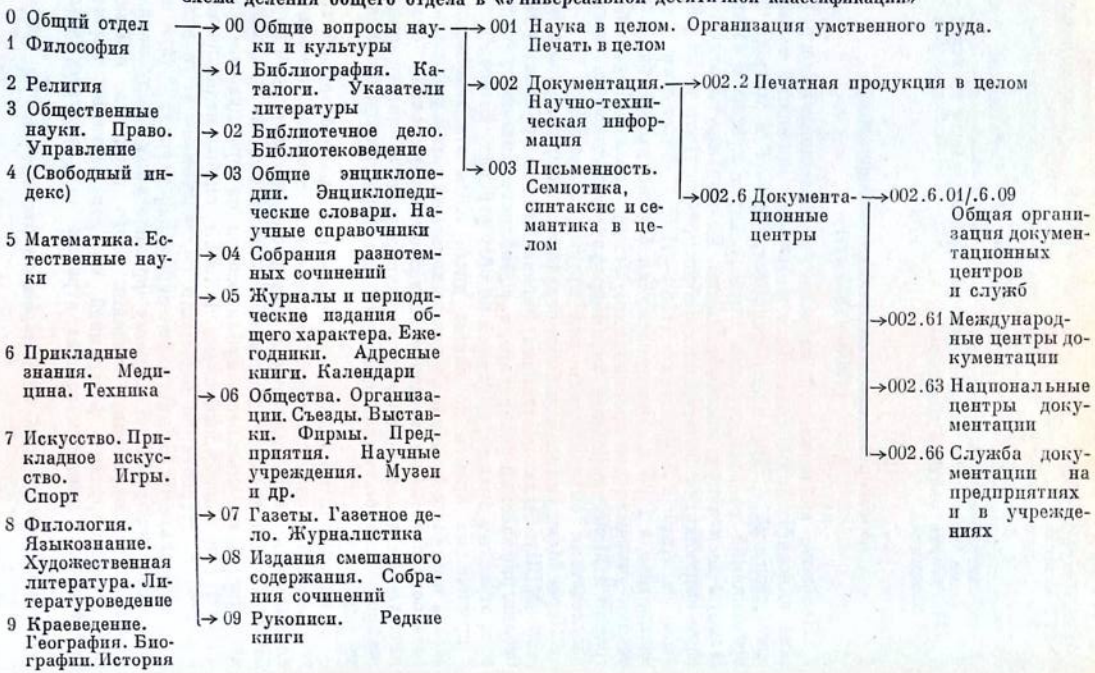
Перед Международным библиографическим институтом в Брюсселе (МБИ) была поставлена задача создать «Универсальный библиографический репертуар», т. е. библиографию литературы по всем отраслям знания, изданной во всех странах, на всех языках с древнейших времен до наших дней. Хотя эта малореальная задача и не была решена, но для выполнения такой работы потребовалось создать детальную классификационную схему, охватывающую все отрасли знания (т. е. универсальную классификацию), которая была бы также пригодна для международного использования.

В основу новой схемы классификации, разработанной группой специалистов во главе с П. Отле, была положена «Десятичная классификация» Дьюи, которая подверглась значительной переработке. Новый вариант десятичной классификации, получивший название «брюссельского», публиковался на французском языке с 1897 г. отдельными выпусками. В 1905 г. он был издан целиком, а в 1907 г. стереотипно переиздан под названием «Руководство к универсальному библиографическому репертуару» [13] (это издание часто цитируют по первому французскому слову заглавия «Манюэль»). Современные издания этого варианта, публикуемые Международной федерацией по документации (преемницей МБИ), а также в разных странах на различных языках,

[13] Manuel du repertoire bibliographique universel. Bruxelles, 1907.

Таблица 6

**Схема деления общего отдела в «Универсальной десятичной классификации»**



получили название «Универсальной десятичной классификации» (УДК).

Основные отделы в УДК по сравнению с «Десятичной классификацией» Дьюи остались неизменными, но их подразделения были более детализированы. Названия некоторых подразделений даны в новых формулировках, изменен внешний вид индексов первой и второй ступени — отброшены нули, стоящие в правой части индексов (табл. 6).

Ниже приводятся основные деления УДК по полным таблицам, изданным в Советском Союзе.

0 Общий отдел
00 Общие вопросы науки и культуры
01 Библиография. Указатели. Каталоги литературы
02 Библиотечное дело. Библиотековедение
03 Общие энциклопедии. Энциклопедические словари. Научные справочники
04 Собрания разнотемных сочинений
05 Журналы и периодические издания общего характера. Ежегодники. Адресные книги. Календари
06 Общества. Организации. Съезды. Выставки. Фирмы. Предприятия. Научные учреждения. Музеи и др.
07 Газеты. Газетное дело. Журналистика
08 Издания смешанного содержания (Собрания сочинений)
09 Рукописи. Редкие книги

1 Философия
10 Общие работы. Философские науки в целом
10 (09) Общая история философии
100 Марксистская философия. Диалектический материализм
11/13 Отдельные философские проблемы
14 Философские системы и направления
15 Психология
16 Логика. Теория познания. Учение о методологии
17 Этика. Учение о морали
18 Эстетика

2 Религия
20 Религия вообще. Сравнительное изучение религий
21 Религия и наука. Вера и знание. Атеизм
22/28 Христианская религия
22 Библия
23 Догматика
24/25 Церковная практика («Практическое богословие»)
26 Устройство церкви и руководство ею
27 Общая история христианской церкви
28 Отдельные христианские церкви, общины и секты
29 Различные религии и культы

3 Общественные науки
30 Социология. Социография
31 Статистика
32 Политические знания. Политика
33. Народное хозяйство. Экономика
34 Юридические науки. Право
35 Управление. Административное право
36 Социальная помощь и обеспечение. Социальное страхование
37 Просвещение. Народное образование. Использование свободного времени
38 Торговля. Транспорт (экономика)
39 Фольклор
4 Отдел 4 освобожден по решению МФД (документ МФД С 42—17 от 19 октября 1962 г.) и его содержание перенесено в отдел 8

5 Математика. Естественные науки
51 Математика
52 Астрономия. Геодезия
53 Физика
54 Химия
55 Геология. Геологические и геофизические науки
56 Палеонтология

57 Биологические науки
58 Ботаника
59 Зоология

6 Прикладные знания. Медицина. Техника
60 Прикладные науки. Общие вопросы
61 Медицина
62 Техника. Инженерное дело
63 Сельское и лесное хозяйство. Охота и рыбоводство
64 Домоводство
65 Руководство производством и организация промышленности, торговли и средств сообщения
66 Химическая технология. Химические производства
67/68 Различные производства
69 Строительные материалы. Строительное дело. Строительные работы.

7 Искусство. Фотография. Зрелища. Спорт
70 Общие работы. Искусство в целом
71 Архитектура городов и населенных мест. Градостроительство
72 Архитектура (отдельных зданий и ансамблей)
73/76 Изобразительные искусства
73 Скульптура и родственные ей виды искусства
74 Рисование и прикладное искусство
75 Живопись
76 Графика. Гравюра
77 Фотография
778. 5 Кинематография
78 Музыка
79 Зрелища. Игры. Спорт

8 Языкознание. Филология. Художественная литература. Литературоведение
80 Языкознание. Филология
82/89 Художественная литература. Литературоведение

9 Краеведение. География. Биографии. История
908 Краеведение
91 География. Географические исследования Земли, отдельных стран и районов
92 Биографии
93/99 История

УДК характеризуется следующими отличительными чертами: 1) охватом всех отраслей знания (в последних ее изданиях содержится более 100 тыс. рубрик); 2) применением десятичного принципа деления, дающего возможность производить неограниченное деление подклассов без нарушения основной структуры классификации; 3) использованием только цифровых индексов, которые сравнительно легко запоминаются и одинаково понятны специалистам, говорящим на разных языках; 4) наличием развитой системы определителей; 5) применением принципа комбинационного построения индексов; 6) возможностью расположения любого числа документов с любой степенью дробности.

Эти черты УДК могут быть проиллюстрированы на примере одного из ее индексов, например 621.311.153.22.

6 Прикладные знания. Медицина. Техника
62 Инженерное дело. Техника
621 Общее машиностроение и машиноведение. Электротехника
621. 3 Электротехника
621.31 Электроэнергетика. Генерирование, преобразование, передача, распределение и регулирование электроэнергии. Электроизмерительная техника. Техническое применение магнетизма и электростатики
621.311 Энергосистемы в целом. Электрические станции и подстанции. Электрификация и сети

621.311.1. Электрификация и проектирование сетей энергоснабжения и электростанций

621.311.15 Мощность станции
621.311.153 Нагрузка
621.311.153.2 Колебания нагрузки
621.311.153.22 Пики нагрузки

Главной отличительной особенностью УДК является наличие в ней развитой системы определителей и применение принципа комбинационного построения индексов. Благодаря этому создание УДК было в известном смысле новым словом в области классификации произведений печати. На смену прежним «перечислительным» схемам с заранее установленными рубриками и готовыми индексами пришла подвижная схема, в которой нужные рубрики создаются в процессе классификации путем или сочетания индексов с определителями, или соединения индексов друг с другом.

В УДК используются общие и специальные (аналитические) определители. К общим определителям относятся: определители языка, формы документов, места, времени, народности, точки зрения. Они используются во всех отделах схемы с одним и тем же неизменным значением. Специальные определители предназначены для использования только в нескольких отделах одной отрасли знания. При помощи таких определителей нижестоящие подразделения этих отделов могут одинаково делиться по одним и тем же признакам.

Так, например, общий определитель -82 может присоединяться ко всем индексам во всех отделах для обозначения произведений, написанных на русском языке; специальный определитель -14 применяется только в подразделении 62 «Техника в целом. Инженерное дело» для обозначения поршневых машин (двигателей) по принципу действия.

По глубине разработки и степени детализации некоторых разделов естествознания и техники УДК представляет собой выдающееся явление в истории библиотечно-библиографической классификации.

Вместе с тем необходимо отметить, что деление всей совокупности человеческих знаний только на десять классов и последовательность этих классов в УДК не соответствуют современному уровню развития науки. Следует особо подчеркнуть, что, поскольку УДК создана буржуазными учеными и специалистами, в ней отразились определенные черты буржуазной идеологии. Это особенно сказалось на таких отделах, как 1 «Философия», 2 «Религия», 3 «Общественные науки». Содержание этих разделов требует пересмотра с позиций марксизма-ленинизма.

Несмотря на эти недостатки, УДК получила широкое распространение, ею пользуются тысячи учреждений более чем в 50 странах мира. В настоящее время объем полных таблиц УДК составляет около 500 авторских листов, т. е. примерно 10 томов по

300—500 страниц в каждом. Работа по переводу и публикации полных таблиц УДК настолько сложна и дорога, что предпринято лишь несколько попыток подобного рода.

Наиболее современными полными изданиями УДК являются: издание на немецком языке (третье международное и первое немецкое, 1934—1953 гг.) [14]; издающееся с 1958 г. и близкое к завершению второе немецкое издание, а также полный перевод разделов 0, 5, 6, 7 на русский язык, опубликованный в 1963—1965 гг. В настоящее время подготовлено второе национальное издание полного перевода УДК на русский язык (за исключением разделов 1 и 2), которое предполагается опубликовать в трех томах в течение 1968—1969 гг.

Кроме полных изданий, широко распространены сокращенные и отраслевые издания УДК; некоторые из них содержат текст одновременно на нескольких языках. В 1958 г., например, Немецким комитетом стандартов, Британским институтом стандартов, Бельгийской ассоциацией по документации и Французским союзом документационных учреждений опубликовано сокращенное издание УДК на английском, немецком и французском языках.

За последние годы под руководством МФД были осуществлены отраслевые издания УДК в Венгрии, Польше, Румынии, Чехословакии, Бельгии, Испании, Италии, Нидерландах, ФРГ, Швейцарии, Японии и других странах. МФД издала отраслевые таблицы УДК по атомной энергии и металлургии. В настоящее время организациями научно-технической информации издано 18 отраслевых таблиц УДК: по химии и химической технологии, телеграфно-телефонной связи и почте, горному делу, сельскому хозяйству, электротехнике, медицине, строительству и архитектуре, машиностроению и др.

Отраслевые таблицы являются наиболее распространенной формой издания таблиц УДК, так как они значительно облегчают пользование Универсальной десятичной классификацией. Преимущество отраслевых схем УДК перед полными таблицами заключается в том, что они содержат все индексы по той или иной отрасли, а смежные разделы включаются лишь в том объеме, в каком это необходимо. В этом случае легче пользоваться и алфавитно-предметным указателем, так как он меньше по объему.

В СССР в 1964—1966 гг. были изданы следующие отраслевые таблицы Универсальной десятичной классификации:

Автомобильный транспорт, автомобильные дороги и смежные области науки и техники. МФД 384. М., «Транспорт», 1965, 112 стр. (НИИАТ).

Геология, геодезия, картография. . . М., «Недра», 1966, 233 стр. (ВИЭМС)

Горное дело. . . М., «Недра», 1964, 128 стр. (ЦНИИТЭИ угля)

Железнодорожный транспорт. . . М., «Транспорт», 1965 (ЦИНТИ ж.-д. транспорта. ЦНТБ МПС СССР).

[14] Dezimalklassifikation. Deutsche Gesamtausgabe. 3. internat. Ausgabe. Berlin, 1934—1953.

Легкая промышленность. . . М., 1964, 233 стр. (ЦНИИТИ легпром)

Лесная, целлюлозно-бумажная, деревообрабатывающая промышленность, лесное хозяйство. . . М., 1964, 152 стр. (ЦНИИТЭИ леспром, ЦНТБ лесной и бумажной промышленности).

Машиностроение. . . М., 1965, 382 стр. (НИИМАШ).

Металлургия. . . М., «Металлургия», 1966, 160 стр. (ЦНИИТЭИ черной металлургии. ЦНИИТЭИ цветной металлургии).

Нефтяная, нефтехимическая, газовая промышленность. . М., 1965, 239 стр. (ЦНИИТЭИ нефтегаз).

Пищевая промышленность. М., 1965. 199 стр. (ЦИНТИ пищепром).

Приборостроение, измерительная техника. М., 1965, 102 стр. (ОНТИПРИБОР).

Проводная телеграфно-телефонная связь, почта. . . М., «Связь», 1965, 79 стр. (ЦНИИсвязи)

Сельское хозяйство. . . М., 1966, 267 стр. (ВИНТИ по сельскому хозяйству).

Строительство и архитектура. . . МФД 397. М., 1965, 162 стр. (ЦИНИ по строительству и архитектуре).

Судостроение. Водный транспорт. . . Л., «Судостроение», 1964, 274 стр. (ГК по судостроению СССР).

Фотография, кинематография. . . М., 1965, 115 стр. (ВНИ кинофотоинститут).

Химия, химическая промышленность. . . М., 1964, 333 стр. (НИИТЭХИМ).

Электротехника. . . Изд. 2 (испр. и доп.). М., 1964, 119 стр. (ВНИИЭМ).

УДК непрерывно пересматривается и обновляется. Эта работа проводится МФД по строго определенным правилам. Любая заинтересованная организация или частное лицо может вносить предложения по изменению того или иного подразделения УДК. После рассмотрения в отраслевых комиссиях и с одобрения Центральной комиссии по классификации МФД эти предложения публикуются в официальном издании МФД «Проекты изменений» («Note-P»). Если на эти предложения не поступит протеста в течение шести месяцев, они считаются принятыми и публикуются в другом официальном издании МФД «Изменения и дополнения к УДК» («Extensions and corrections of the UDC»), а также в издании комиссии по классификации Германского комитета стандартов «ДК-сообщения» («DK-Mitteilungen»), в котором перепечатываются целиком страницы третьего международного (немецкого) издания УДК с соответствующими исправлениями. Из крупных изменений УДК в последние годы следует назвать освобождение 4-го отдела «Филология», материал которого перенесен в 8-й отдел «Литература» (октябрь 1962 г.). По существующим правилам 4-й отдел в течение десяти лет должен оставаться свободным, после чего его можно будет заполнить новым содержанием.

За последние годы серьезной переработке подверглись такие разделы УДК, как ядерная энергетика, металлургия, каучук и пластмассы, текстиль, география, геодезия, авиация и космонавтика, вычислительная техника и др. Особенно интенсивно сейчас ведется пересмотр следующих разделов УДК: астрономия,

математика, науки о Земле, радиотехника, горное дело, социальные науки, химия и химическая технология и др.

Наличие хорошо организованной международной системы для поддержания УДК на современном уровне является ее несомненным достоинством, выгодно отличающим УДК от любой другой из существующих классификаций. Естественно, что создание такой системы требует длительного времени, а ее функционирование связано со значительными трудностями и затратами. Однако только при ее наличии возможно постоянное развитие и совершенствование любой схемы классификации.

Наряду с текущей работой по дополнению и изменению УДК, в настоящее время МФД предпринята серьезная работа по радикальному улучшению УДК, которая связана с исследованиями по использованию УДК для механизированного поиска. В МФД создан специальный комитет по механизации УДК, в который вошли крупные специалисты многих стран.

Было бы неправильно отрицать большие достоинства УДК: универсальность, международную применимость, десятичную систему индексации, принцип комбинационного построения индексов, налаженную организацию для поддержания ее на современном уровне. На данном этапе развития классификационной теории и практики УДК остается единственной международной универсальной системой, позволяющей достаточно детально раскрывать содержание справочно-информационных фондов, обеспечить быстрый поиск информации, усилить международное сотрудничество [15].

Эти достоинства УДК делают безусловно целесообразным ее использование в советских органах научно-технической информации и специальных библиотеках для классификации естественнонаучной и технической литературы. Постановлением Совета Министров СССР от 11.V 1962 г. УДК введена для органов научной информации и технических библиотек как обязательная система классификации литературы по естественным и техническим наукам.

## «Классификация с двоеточием» Ш. Р. Ранганатана (1933 г.)

Шиали Рамамрита Ранганатан (род. в 1892 г.) является одним из наиболее оригинальных современных библиотековедов, работающих в области теории классификации. Перу этого выдающегося индийского ученого и общественного деятеля принадлежат многие фундаментальные работы, в которых обнаруживается стройная система взглядов на библиотечное дело, библиографию,

[15] Более подробные сведения о построении и использовании таблиц УДК приведены ниже при изложении методики классификации.

документацию. Они служат последовательным развитием принципов, сформулированных в его первых книгах «Пять законов библиотечной науки» (1931 г.) [16] и «Классификация с двоеточием» (1933 г.) [17].

Сформулированные им в первой книге законы: 1) «книги — для пользования», 2) «каждому читателю — его книгу», 3) «каждой книге — ее читателя», 4) «береги время читателя», 5) «библиотека — растущий организм» — полемизируют с господствующими в зарубежных научных библиотеках принципами, по которым книги собираются в библиотеке для сохранения и использования немногими избранными. Ранганатан критикует классовый характер библиотечного дела в буржуазном обществе и связывает необходимость перестройки работы библиотек с культурной революцией.

Книга, в которой Ранганатан изложил свою систему и взгляды на теорию классификации, — «Классификация с двоеточием» получила широкую известность. Она неоднократно переиздавалась с изменениями и дополнениями (в 1939, 1950, 1952, 1957 и 1960 гг.). Новизна предложенной Ранганатаном системы заключалась в том, что он выступил против практики составления подробных «перечисляющих» классификационных таблиц, в которых составители стремятся предусмотреть самостоятельные индексы для каждого предмета и понятия. Это, по его мнению, делает классификацию громоздкой и негибкой, не позволяет достаточно подробно анализировать каждый документ при его классификации.

Развивая идею, заложенную П. Отле и А. Лафонтеном в «Универсальной десятичной классификации», Ранганатан создает свою систему на основе «комбинационного» принципа. Таблица основных делений его классификации насчитывает всего 25 классов, последовательность которых взята из философских и религиозных толкований древнеиндийских священных книг. Частично она совпадает с традиционной группировкой знаний, встречающейся в других библиотечных классификациях XIX и XX вв.

*A* Наука (вообще)
*B* Математика
*C* Физика
*D* Инженерное дело
*E* Химия
*G* Естественные науки (вообще) и биология
*H* Геология
*I* Ботаника
*J* Сельское хозяйство
*K* Зоология
*L* Медицина
*M* Полезные искусства
*N* Изящные искусства
*O* Литература
*P* Языковедение
*Q* Религия
*R* Философия
*S* Психология
*T* Воспитание
*U* География
*V* История
*W* Политика
*X* Экономика
*Y* Смешанные материалы по общественным наукам
*Z* Право

[16] S. R. Ranganathan. The five laws of library science. Madras—London, 1931.
[17] S. R. Ranganathan. Colon classification. Madras, 1933.

Вместо единого ряда делений в каждом основном классе в системе Ранганатана разработаны таблицы, каждая из которых построена на основе какой-либо одной характеристики или аспекта, получивших в дальнейшем название «фасетов». Индекс, отражающий содержание каждого документа, строится из обозначений, принятых в каждой такой таблице и соединяемых при помощи двоеточия. Этот метод комбинации индексов и дал название самой системе классификации.

«Термин „фасет" Ранганатан понимает как аспект, в котором предмет может рассматриваться внутри определенной отрасли знания или науки. Так, в медицине определенное понятие может рассматриваться с точки зрения органа или системы органов, с точки зрения характера заболевания, с точки зрения методов лечения и т. д. В строительном деле отдельными „фасетами" могут быть вид или назначение постройки, материал, из которого она построена, процессы строительных работ, место, где она построена (страна, населенный пуект), время постройки. Каждый из этих аспектов представляет собой „фасет" рассмотрения предмета и для каждого из них предусматривается специальная стандартная табличка делений. Подразделение „фасета", то есть единичное деление стандартной таблички, Ранганатан называет „фокусом", а процесс определения соответствующего деления, адекватного содержанию классифицируемой книги или статьи, „фокусированием". Процесс этот, по его словам, сводится к снижению экстенсивности и увеличению интенсивности понятия внутри одного „фасета" с целью достичь объема специфического деления, то есть „фокуса". Под термином „фасетная классификация" Ранганатан понимает „полное поочередное исчерпывание каждой характеристики и отражение в индексе каждого последовательного фасета". Систему индексации, при которой в индексе находит свое отражение каждый „фасет", он называет „фасетной индексацией".

При классификации материалов по системе Ранганатана важнейшее значение приобретает правильное определение „фасетов", которые отражают содержание книги, и правильное установление последовательности их применения. Это достигается тщательно проводимым „фасетным анализом" материала и выведенной на основе такого анализа стандартной „фасетной формулой". Как все многообразие цветов, существующих в природе, может быть сведено к сочетаниям четырех основных, так и все разнообразные „характеристики", то есть основания деления, которые используются при классификации самых разнообразных предметов и явлений, базируются на свойствах или качествах последних, сводимых к пяти основным „категориям": Персональность („индивидуальность"), Материя, Энергия, Пространство и Время. Четыре последних являются „основными понятиями для всех наук и отраслей знания", а первое — это качество, присущее „целост-

ности" каждой вещи, свойство, обусловливающее различие и многообразие существующих в природе предметов и явлений. Составление классификаций чисто эмпирическим путем, без предварительно продуманных „фасетов" и намеченной на основе „фасетного анализа" „фасетной формулы", приводило, по мнению Ранганатана, к случайности и разнородности построения даже родственных отделов, что наблюдается, например, в „Универсальной десятичной классификации". Применение „фасетного анализа" обеспечивает, напротив, плановость как разработки классификации, так и процесса классификации. Сам Ранганатан принял такую последовательность „фасетов", то есть „фасетную формулу", делая оговорку, что он не настаивает на безусловной ее правильности и надеется, что будущие исследователи внесут в нее поправки и дополнения:

Индивидуальность (класс, подкласс или подразделение по предмету),

Материя (материал),

Энергия (операция, процесс, действие),

Пространство (географический момент, территория),

Время» [18].

Для пояснения довольно сложной методики индексации по системе Ранганатана приводим пример из области медицины (индекс L).

| Фасет «О» | Фасет «Р» | Фасет «Н» |
|---|---|---|
| Органы человеческого тела | Проблемы медицины | Уход и лечение |
| 1 Организм в целом | 1 Общие проблемы | 1 Питание |
| 2 Органы пищеварения | 2 Морфология | 2 Этиология |
| 23 Пищевод | 3 Физиология | 3 Диагностика |
| 24 Желудок | 4 Болезни | 4 Лечение |
| 25 Кишечник | 41 Общие болезни | |
| 3 Кровеносная система | 411 Атрофия | |
| 32 Сердце | 42 Инфекционные болезни | |
| 35 Кровь | 421 Туберкулез | |
| 36 Артерии | | |
| 4 Органы дыхания | | |
| 45 Легкие | | |

При классификации документов по диагностике инфекционных заболеваний кишечника они получат индекс: L 25 : 42 : 3; по лечению туберкулеза легких: L 45 : 421 : 4.

Работа Ранганатана в области теории классификации не ограничивается совершенствованием его системы. Он считает необходимым создание еще более дробной, конкретной и углубленной классификации, которая удовлетворяла бы потребности массового читателя в демократизированной библиотеке и отвечала совре-

[18] Е. И. Шамурин. К семидесятилетию Ш. Р. Ранганатана. «Сов. библиография», 1962, № 5, стр. 105—106.

менному уровню развития науки и техники. Эти взгляды нашли отражение в книгах «Пролегомены в библиотечной классификации» (1937; изд. 2, 1957), «Библиотечная классификация: основные принципы и процесс» (1950) и других работах.

Большой интерес представляет предложенный Ранганатаном способ расширения базы индексации в десятичной системе, а также «цепной метод» составления предметного указателя к систематическому каталогу. Как известно, на каждой ступени деления десятичной системы всего десять или даже девять делений, так как 0 везде сохраняется как символ «общего деления». Ранганатан предлагает использовать только восемь делений (от 1 до 8), а 9 — для дальнейшего подразделения той же ступени. Таким образом, любая ступень деления может содержать неограниченное число делений без нарушения структуры десятичной системы, например:

I ступень: 1, 2 ... 8, 91, 92 ... 98, 991, 992 ... 998 ...
II ступень: 11, 12 ... 18, 191, 192 ... 198, 1991, 1992 ... 1998 ...
III ступень: 111, 112 ... 118, 1191, 1192 ... 1198, 11991, 11992 ... 11998 ...

Этот способ, получивший название «октавы», признан исключительно плодотворным и в 1948 г. принят для нового издания УДК.

Сущность «цепного метода» заключается в том, что в указатель выносятся термины, отражающие каждую ступень индекса, т. е. логически соподчиненные понятия всех ступеней деления. «Цепной метод» можно проиллюстрировать примером, приведенным выше. Индекс L 45 : 421 : 4 «Лечение туберкулеза легких» получает в указателе следующие рубрики:

Лечение туберкулеза легких L 45 : 421 : 4
Туберкулез легких L 45 : 421
Инфекционные болезни легких L 45 : 42
Болезни легких L 45 : 4
Легкие (медицина) L 45
Органы дыхания (медицина) L 4
Медицина L

Классификационная система Ранганатана не находит широкого практического применения даже на родине ее автора, в Индии, тем не менее значение ее исключительно велико. Использовав все наиболее ценное в теории классификации, что было сделано его предшественниками, Ранганатан выдвинул принципы, обусловившие переход от «перечисляющих» классификаций к «комбинационным». Антинаучность расположения знаний в основных делениях, сложность индексации, неприменимость единой фасетной формулы членения для различных отраслей знания — все эти слабые стороны «Классификации с двоеточием» обусловливают нецелесообразность ее практического применения.

Достоинства системы Ранганатана и его классификационных идей — принцип фасетного анализа отраслей знания, большие комбинационные возможности «Классификации с двоеточием», ее неограниченная растяжимость и разработанность правил классификации — обеспечили им большую популярность во всех странах мира. Классификационная теория Ранганатана положена в основу почти всех современных разработок в этой области, она дает ценный отправной материал при создании информационно-поисковых языков дескрипторного типа.

### «Библиографическая классификация» Г. И. Блисса (1935 г.)

Автор этой системы — американский библиотековед Генри Ивлин Блисс (1870—1955 гг.) — строит ее на основе собственной классификации наук. Особенностью его системы является то, что каждую отрасль знания Блисс преломляет в четырех аспектах:

1) философия данной науки,
2) собственно наука,
3) история науки,
4) практическое применение науки.

При этом все четыре аспекта объединяются в одном классе. Таким образом, класс химии включает не только теоретическую химию, но и химическую технологию и промышленность (табл. 7).

Система Блисса содержит несколько таблиц. Рубрики строятся путем комбинирования обозначений, взятых из различных таблиц в соответствии с содержанием и формой документов. Вспомогательные таблицы типовых делений подразделяются на три группы:

I. Таблицы делений, применяемых во всех разделах, — общие (формы) и географические.

II. Исторические и филологические таблицы, применяемые в разделах общественных и гуманитарных наук.

III. Специальные таблицы, применяемые в одном каком-либо разделе (химия, астрономия, ботаника, религия и т. д.).

Таблица основных классов содержит деления двух видов. Первые девять классов с цифровыми обозначениями предназначаются для разных частей книжного фонда [19].

1. Фонды читального зала, преимущественно справочные.
2. Книговедение, библиография и библиотечное дело (альтернатива к классу Z).
3. Избранные или специальные фонды, а также книги, выделенные в отдельное помещение.
4. Фонды отделов и специальные фонды.

---

[19] H. E. Bliss. A bibliographic classification. Extended by systematic auxilliary schedules for composite specification and notation. 2nd ed. V. 1—4. New York, 1952—1953.

Таблица 7

**Краткая синоптическая таблица последовательности классов в «Библиографической классификации» Блисса**

| Философия | Наука | История | Техника и искусства |
|---|---|---|---|
| Основные принципы<br>Метафизика | Наука в целом<br>*Абстрактные науки и общие принципы*<br>Математика | История философии и наука | |
| Философия природы | *Естественные науки*<br>*Физические науки*<br>Физика<br>Химия<br>*Специальные естественные науки и описательная естественная история* | | *Прикладные науки*<br>Технология:<br>физическая;<br>техническая;<br>химическая |
| Космология | Астрономия<br>Геология<br>География<br>Метеорология | Историческая геология | Экономическая геология<br>Экономическая география |
| Философия жизни | *Биологические науки*<br>Биология<br>Ботаника<br>Зоология | | |
| Философия человеческой жизни | *Антропологические науки*<br>Антропология | *Гуманитарные науки*<br>История человечества и человеческой жизни<br>Археология | |
| Философия духа<br>Философия общества | *Психологические науки*<br>Психология<br>Социальная психология<br>*Общественные науки*<br>Социология<br>Этнология | Общественно-политическая история | Медицина и гигиена<br>Прикладная психология<br>Психиатрия<br>Воспитание<br>Прикладные науки<br>Общественные науки |
| Теология<br>Этика<br>Политическая экономия | Религия<br>Мораль<br>Политические науки<br>Юриспруденция<br>Экономика | История религий, мифология<br>История права<br>История экономики | Церковная деятельность<br>Прикладная этика<br>Управление государством<br>Юридическая практика<br>Хозяйство<br>Промышленность<br>Торговля |
| Эстетика | Филология<br>Лингвистика и языки | Литература | Искусства:<br>История искусств:<br>Техника искусств:<br>прикладных;<br>изящных и т. п.;<br>музыки<br>Риторика, ораторское искусство и т. п.<br>Драматургия и театр |

5. Документы или архивы правительства, учреждений и т. д.
6. Периодические издания.
7. Смешанные материалы.
8. Собрания материалов исторического или местного значения.
9. Антикварные книги или собрания, представляющие исторический интерес.

Остальные главные классы основной таблицы с буквенными индексами охватывают различные отрасли знаний. При этом последовательность наук следует порядку эволюционного развития мира: физико-химические науки, науки о неорганической природе, биологические науки, науки о человеке и обществе.

Основные классы

A Философия и наука в целом,
включая логику, математику, метрологию и статистику.
B Физика,
включая прикладную физику и физическую технологию.
C Химия,
включая химическую технологию и промышленность, а также минералогию.
D Астрономия, геология, география и естественная история,
включая микроскопию. *География имеется в виду только общая и физическая.*
E Биология,
включая палеонтологию и биогеографию.
F Ботаника,
включая бактериологию.
G Зоология,
включая зоогеографию и экономическую зоологию.
H Антропология, общая и физическая,
включая медицинские науки, гигиену, евгенику, физическую культуру, игры и спорт.
I Психология,
включая сравнительную психологию, расовую психологию и психиатрию. *Альтернативно отражается в классе «Философия».*
J Воспитание,
включая психологию воспитания.
K Общественные науки:
социология, этиология и антропогеография. *Альтернативно здесь может отражаться класс P «Религия и Этика».*
L История социальная, политическая и экономическая,
включая историческую и политическую географию и этнографию, нумизматику и другие вспомогательные дисциплины. Древняя история, средневековая история и общая история нового времени.
M Европа.
N Америка.
O Австралия, Азия, Африка и Океания.
География, этнография и история.
P Религия, богословие и этика. Альтернативно отражается в классах A, K или R.
Q Прикладные общественные науки и прикладная этика.
R Политические науки, философия, этика и политическая деятельность.
S Юриспруденция и право.
T Экономика.
U Прикладное искусство, ремесла и промышленная технология.

V Изящные искусства и искусства выражения, развлечения и времяпрепровождения.
W Филология: лингвистика и неиндоевропейские языки.
X Индоевропейская филология, языки и литературы.
Y Английский (или какой-либо другой) язык и литература, а также литература в целом, риторика, ораторское искусство, драматургия.
Z Книговедение, библиография и библиотечное дело. *Альтернативно отражается в классе 2.*

Система Блисса исходит из принципа альтернативности. Это значит, что схема предусматривает для некоторых вопросов несколько различных мест, которые выбираются в зависимости от конкретных условий библиотеки. Примерами альтернативных решений служит отдел психологии, который может выделяться самостоятельно или включаться в отдел философии, а также отдел техники, который можно размещать в подразделениях физики или в отделе прикладных наук. Широкое применение этого принципа лишает схему Блисса устойчивости — некоторые важные отделы не имеют в ней постоянного места.

В «Библиографической классификации» отражается реакционная идеология господствующих классов США: этика включена в отдел религии, биологические законы естественного отбора и борьбы за существование рассматриваются как факторы общественного развития, рубрика «Революция» помещена рядом с рубрикой «Социальная дезорганизация» и т. п. Все эти и многие другие недостатки «Библиографической классификации» не позволяют нам присоединиться к восторженным оценкам системы Блисса со стороны многих буржуазных библиотековедов. Тот факт, что это — одна из подробно разработанных современных схем с некоторыми элементами комбинационного построения индексов, еще не говорит о ее совершенстве. Даже среди традиционных линейно-иерархических классификационных схем эта система не является лучшей. Об этом свидетельствует и тот факт, что практически она применяется очень незначительно.

### *Советские схемы классификации*

В послеоктябрьские годы в практику советских библиотек стала внедряться десятичная классификация. Несмотря на ее недостатки и идеологические пороки, это была самая авторитетная, детально разработанная и проверенная опытом система классификации. В 1921 г. в целях унификации классификационной практики в стране Главполитпросвет РСФСР принял постановление «Об обязательном переводе всех библиотек РСФСР на Брюссельскую международную систему десятичной классификации». Для реализации этого постановления в том же году были изданы сокращенные таблицы десятичной классификации.

Однако в этих таблицах сохранились идеологические недостатки УДК, не были в ней учтены многие новые явления советской действительности. Отдельные исправления в переизданиях этих таблиц 1927 и 1931 гг. не внесли сколько-нибудь существенных изменений. Стало ясно, что десятичная классификация нуждается в серьезной переработке.

Эта задача была выполнена видным советским библиографом Л. Н. Троповским, таблицы которого вышли в 1938 г. В этих таблицах был образован отдел «Марксизм-ленинизм», помещенный перед отделом общественно-политической литературы, предусмотрено широкое отражение во всех отделах произведений основоположников марксизма-ленинизма, детализированы отделы для документов, посвященных деятельности КПСС, ВЛКСМ, коммунистических партий и союзов молодежи зарубежных стран, разграничена литература, освещающая общественный, государственный и экономический строй СССР, с одной стороны, и капиталистических стран, — с другой, изменена формулировка многих рубрик в соответствии с советской научной терминологией.

В связи со всеми этими изменениями потребовалось нарушить структуру индексов десятичной классификации, внеся в нее буквенные обозначения. Таблицы Л. Н. Троповского получили широкое распространение в массовых библиотеках страны. Применяемые в настоящее время «Таблицы библиотечной классификации для массовых библиотек» (изд. 2. М., 1961) и «Таблицы библиотечной классификации для областных библиотек» (М., 1963), изданные под редакцией З. Н. Амбарцумяна, являются дальнейшей переработкой таблиц Л. Н. Троповского.

Научные библиотеки, для которых таблицы Л. Н. Троповского оказались недостаточно детализированными, использовали более подробные переработки десятичной классификации, из которых наибольшей известностью пользовались последние издания таблиц Е. Н. Добржинского (1931 г.) и Н. В. Русинова (1944 г.). Они применялись главным образом для классификации естественно-научной и технической литературы в крупных специальных библиотеках.

Как уже было отмечено, в настоящее время органы научно-технической информации, технические библиотеки, редакции журналов и издательства, выпускающие соответствующую литературу, начали с 1963 г. применять Универсальную десятичную классификацию, что потребовало издания таблиц общих определителей и классов 0; 5; 6; 7. В большинстве изданий текущей регистрационной библиографии литература классифицирована по схеме Всесоюзной книжной палаты, состоящей из 31 отдела. Индексы в этой схеме смешанные, буквенно-цифровые, ступенчатые, в них используются римские и арабские цифры и строчные буквы русского алфавита.

Е7 Антропология
Е8 Эмбриология, анатомия и гистология человека
Е9 Физиология, биофизика и биохимия животных и человека

Ж/О Техника. Технические науки
Ж Техника и технические науки в целом
З Энергетика. Радиоэлектроника
И Горное дело
К Технология металлов. Машиностроение. Приборостроение
Л Химическая технология, химические производства. Пищевые производства
М Технология древесины. Производства легкой промышленности. Полиграфическое производство. Фотокинотехника
Н Строительство
О Транспорт
П Сельское и лесное хозяйство. Сельскохозяйственные и лесохозяйственные науки
П0 Естественнонаучные и технические основы сельского хозяйства
П1/2 Растениеводство
П3 Лесное хозяйство
П4 Защита растений
П5/6 Животноводство
П7 Охотничье хозяйство. Рыбное хозяйство
П8 Ветеринария
П9 Сельское и лесное хозяйство отдельных территорий
Р Здравоохранение. Медицинские науки
Р11 Организация здравоохранения
Р12 Гигиена
Р19 Эпидемиология
Р25 Общая патология
Р26 Медицинская микробиология и паразитология
Р28 Фармакология. Фармация. Токсикология
Р34 Общая диагностика
Р35 Общая терапия
Р36 Медицинская радиология и рентгенология
Р41 Внутренние болезни
Р45 Хирургия
Р51 Инфекционные и паразитарные болезни
Р54 Фтизиология
Р56 Онкология
Р58/60 Венерология. Дерматология
Р61/64 Невропатология. Нейрохирургия. Психиатрия
Р66 Стоматология
Р67 Офтальмология
Р68 Оториноларингология
Р69 Урология
Р71 Гинекология
Р73 Педиатрия
Р81 Судебная медицина
Р86 Авиационная медицина
Р87 Космическая медицина
Р89 Военная медицина
С Общественные науки в целом
С6 Статистика
С7 Демография
Т История. Исторические науки
Т0 Теоретические основы и методология исторической науки
Т1 История исторический науки
Т2 Источниковедение. Вспомогательные исторические дисциплины
Т3 История
Т4 Археология
Т5 Этнография
У* Экономика. Экономические науки
Ф* Коммунистические и рабочие партии. Общественно-политические организации трудящихся
Х Государство и право. Юридические науки
Х0 Общая теория государства и права
Х1 История политических учений
Х2 История государства и права
Х6 Государство и право социалистических стран
Х7 Государство и право стран, освободившихся от колониальной зависимости
Х8 Государство и право капиталистических стран
Х91 Международное право (международное публичное право)
Х93 Международное частное право
Х99 Отрасли права
Ц Военная наука. Военное дело
Ц0 Учение о войне и армии
Ц1 Общая теория военной науки и история военно-теоретической мысли
Ц2 Теория военного искусства

Ц3 Военно-историческая наука
Ц4/7 Вооруженные силы
Ц8 Военная география
Ц9 Военная техника. Военно-технические науки

Ч Культура. Наука. Просвещение
Ч11 Культура. Культурное строительство
Ч21 Наука. Научно-исследовательская работа
Ч23 Научная информация
Ч30/49 Народное образование. Педагогические науки
Ч51 Физическая культура и спорт
Ч61 Печать
Ч63 Радиовещание
Ч71 Культурно-просветительная работа
Ч72 Клубное дело
Ч73 Библиотечное дело. Библиотековедение
Ч75 Библиография
Ч77 Музейное дело. Музееведение
Ч81 Архивное дело. Архивоведение

Ш Филологические науки. Художественная литература
Ш0 Частные филологии. Текстология
Ш1 Языкознание
Ш2/3 Фольклор (народно-поэтическое творчество). Фольклористика
Ш2/6 Художественная литература. Литературоведение
Ш7 Ораторское искусство
Ш8/9 Детская литература

Щ Искусство. Искусствознание.
Щ03 История искусства
Щ10 Изобразительное искусство
Щ11 Архитектура
Щ12 Декоративно-прикладное искусство
Щ13 Скульптура
Щ14 Живопись
Щ15 Графика
Щ16 Художественная фотография
Щ30 Музыка и зрелищные искусства
Щ31 Музыка
Щ32 Танец (хореография)
Щ33 Театр
Щ34 Массовые представления и театрализованные праздники
Щ35 Цирк
Щ36 Эстрада
Щ37 Киноискусство
Щ38 Художественное радиовещание и художественное телевидение
Щ7 Художественная самодеятельность
Щ8 Изобразительные материалы (печатная графика)
Щ9 Музыкальные произведения (нотные издания)

Э Религия. Атеизм
Э0 Марксизм-ленинизм о религии и атеизме
Э1 Свободомыслие и атеизм
Э2/9 Религия

Ю* Философские науки. Психология

Я Литература универсального содержания
Я1 Библиографические указатели
Я2 Справочные издания
Я4 Серии. Сборники. Собрания сочинений
Я5 Периодические и продолжающиеся издания
Я6 Иллюстрированные издания и материалы
Я9 Популярные издания

В новой советской схеме широко применяются приемы, обусловившие успех десятичной классификации, — ступенчатая десятичная структура индексов для подразделений каждого отдела, комбинационный принцип построения индексов, детально разработанные таблицы типовых делений. Вместе с тем примененная в схеме буквенно-цифровая система индексов с использованием различных знаков препинания представляется усложненной по сравнению с УДК.

### *Методика классификации*

Методика классификации состоит из совокупности методов и практических приемов распределения документов в соответствии с их содержанием по отделам и подразделениям принятой схемы классификации. Она распадается на общую методику, в которой рассматриваются методы и приемы классификации документов любого содержания, независимо от отрасли знания и формы документа и частную методику, касающуюся лишь документов какой-либо определенной отрасли знания или конкретного вида [21].

Процесс классификации распадается на три последовательных этапа:

1) анализ документа, в результате чего выявляется его основное содержание, информационная и научная ценность, вид, читательское назначение и все аспекты, в которых рассматриваются отраженные в нем вопросы;

2) решение, т. е. определение тех аспектов содержания и формы документа, по которым целесообразно составить его поисковые образы, а при систематической расстановке самих документов — установление главного аспекта, определяющего место данного документа среди других;

3) индексация или запись принятого решения в условных обозначениях (индексах) той или иной схемы классификации. Этот последний этап предполагает знание структуры и конкретных особенностей самой схемы. В настоящем издании методика классификации излагается применительно к Универсальной десятичной классификации, принятой во всех советских органах научно-технической информации и технических библиотеках.

Как было уже сказано, Универсальная десятичная классификация состоит из основной таблицы, вспомогательных таблиц и алфавитно-предметного указателя. Основная таблица содержит индексы, при помощи которых систематизируются документы по содержанию. Каждому понятию, отраженному в основной таблице, должен соответствовать определенный индекс УДК. Вся сумма человеческих знаний делится на 10 основных отделов:

0 Общий отдел
1 Философия
2 Религия
3 Общественные науки
4 (Отдел свободен до октября 1972 г.)
5 Математика. Естественные науки.
6 Прикладные знания. Медицина. Техника
7 Искусство. Фотография. Зрелища. Спорт.
8 Языкознание. Филология. Художественная литература. Литературоведение
9 Краеведение. География. Биографии. История

---

[21] П. Герман. Практическое применение Универсальной десятичной классификации. Техника классификации. Пер. с нем., изд. 5, доп. и переработ. М., 1967.

Каждая последующая цифра, добавляемая к обозначению основного отдела, не меняет общего содержания индекса, а лишь уточняет его содержание. Подобное построение индексов УДК позволяет делить каждое самое общее понятие на узко специальные вопросы. Чем больше степень дробления понятия, тем большее число знаков входит в индекс. Так, например, прочность грунтов на сжатие в строительной технике обозначается индексом:

624.131.439.4

Для лучшей обозримости индекса после каждой третьей его цифры ставится точка. Индексы произносятся как комплексы целых чисел (шестьсот двадцать четыре, сто тридцать один, четыреста тридцать девять, четыре).

Индексы УДК располагаются по принципу от общего к частному в пределах одного раздела. Их последовательность зависит только от порядка цифр в каждом разделе, а не от длины индекса.

Наряду с основными индексами в основной таблице приведены специальные (аналитические) определители, отражающие признаки и свойства, характерные для узкого круга понятий и применяемые лишь в том разделе, в начале которого они помещены. В большинстве изданий УДК аналитические определители отчеркнуты вертикальной линией. Они присоединяются к основным индексам при помощи дефиса (определители, применимые во многих подразделениях данного отдела) или при помощи точки и нуля (определители, применимые только в данном подразделении). Так, например, в разделе 62 «Инженерное дело. Техника» применяются специальные определители:

62-1 Общая характеристика машин, аппаратов и установок
-2 Подвижные и неподвижные детали машин, аппаратов и установок
-3 Детали распределительных механизмов. Приводы распределительных механизмов
-4 Состояние (агрегатное) материала, состояние поверхности, форма и конфигурация предметов
-5 Работа машин, установок, аппаратов. Регулирование и управление
-6 Характеристика двигателей по виду топлива
-7 Обслуживание машин, аппаратов и установок. Уход за ними
-8 Характеристика машин, аппаратов и установок по виду привода, по роду используемой энергии
-97/98 Рабочие и эксплуатационные характеристики

Пример специальных (аналитических) определителей с «точкой — нулем» может быть взят из раздела 621.3 «Электротехника»:

621.3.01 Теория и основные понятия
.02 Виды и диапазоны токов, мощностей, напряжений, сопротивлений и частот
.03 Снабжение электроэнергией в специальных цехах. Части и детали электрических и электронных ламп, электронно-лучевых и газоразрядных трубок, электрохимического оборудования и аппаратуры, электронного оборудования и электрических печей

.04 Детали, изоляция и обмотки электрических машин, трансформаторов и аппаратов. Монтаж схем
.05 Передача электроэнергии. Система передачи. Сооружение сетей
.06 Соединение и коммутация электрических цепей. Включение и отключение
.07 Электрическое регулирование
.08 Теория и методы электрических измерений. Конструктивные элементы и узлы электроизмерительных приборов
.09 Характеристики передач

Во вспомогательных таблицах УДК перечислены общие определители, отражающие признаки, по которым документы могут подразделяться во всех отраслях знаний, и применимые во всех отделах УДК, а также различные знаки, служащие для соединения нескольких основных индексов УДК (табл. 9).

Таблица 9

**Общие определители УДК**

| Название | Обозначение | Произношение | Примеры |
|---|---|---|---|
| Определитель языка . . . | = | Равенство | =30 на немецком языке |
| Определитель формы и характера материала . . . . | (0) | Скобки, нуль | (083.57) нормативы |
| Определитель места . . . . | ( ) | Скобки | (47+57) Советский Союз |
| Определитель рас и народов . . . . . | (=) | Скобки, равенство | (=927) арабы |
| Определитель времени . . | « » | Кавычки | «1964.01.24» 24 января 1964 г. |
| Определитель точки зрения | .00 | Точка, нуль, нуль | .003 Экономическая точка зрения |

К индексам УДК в случае необходимости могут добавляться буквы латинского или русского алфавитов, отдельные слова или имена, например:

92 Эйнштейн (Биография А. Эйнштейна).

Для более полного и точного отражения содержания классифицируемых документов в УДК, наряду с основными индексами и определителями, применяется соединение нескольких индексов при помощи различных знаков (табл. 10).

Знак присоединения применяется в тех случаях, когда в документе рассматривается несколько самостоятельных вопросов

Таблица 10

**Виды соединений индексов УДК**

| Название | Обозначение | Произношение | Примеры |
|---|---|---|---|
| Присоединение . . . . . | + | И | 5+62 естествознание и техника |
| Распространение . . . . . | / | От, до | 546.32/.34 от калия до лития |
| Отношение . . . . . | : | Двоеточие | 016 : 621.3 библиография по электротехнике |
| Объединение (синтез) . . . . | ' | Апостроф | 546.32'175 нитрат калия |

равного значения. Знак распространения служит для замены индексов, отражающих близкие по значению вопросы и последовательно расположенных в таблице. Знак отношения позволяет классифицировать документы, трактующие о вопросах, между которыми имеется внутренняя связь. Знак объединения (синтеза) употребляется при классификации материалов по химическим соединениям или сплавам.

Поскольку соединение нескольких индексов и применение определителей осуществляется самим систематизатором и в таблицах УДК порядок расположения различных видов индексов предусмотрен быть не может, в общей методике классификации принята определенная последовательность составных индексов УДК (табл. 11).

Классификация по УДК значительно облегчается благодаря наличию в таблицах алфавитно-предметного указателя. Установив основное содержание документа, необходимо обратиться к алфавитно-предметному указателю и определить по нему раздел или подразделение УДК, соответствующее содержанию документа. Затем следует познакомиться с содержанием соответствующего раздела или подраздела, выбрать наиболее подходящую рубрику и записать ее индекс. При определении рубрики не обязательно следовать последней степени дробности применяемых таблиц. В этом случае целесообразно руководствоваться наличием материала по данному вопросу и перспективой увеличения документов аналогичного содержания. Однако применение укрупненных индексов нежелательно, так как при этом теряется глубина классификационного анализа документа.

Следует учитывать, что индексирование по УДК не позволяет отразить всех аспектов, в которых рассматривается данный вопрос в документе. Классификационный индекс не может заменить аннотации. Поэтому из всех аспектов содержания документа при классификации выбирают только основные, но их желательно отразить

Таблица 11

Расположение составных индексов УДК *

| Виды соединений и определители | Индекс | Содержание индекса |
|---|---|---|
| Распространение . . . . | 621.385.2/.5 | От диодов до пентодов |
| Простой индекс . . . . | 621.385.2 | Ламповые диоды |
| Общий определитель языка . . . . . | 621.385.2-30 | Литература по ламповым диодам на немецком языке |
| Общий определитель формы и характера материала . . . . . | 621.385.2(031) | Справочник по ламповым диодам |
| Общий определитель места . . . . . | 621.385.2 (73) | Ламповые диоды США |
| Общий определитель времени . . . . . | 621.385.2 «1960» | Ламповые диоды 1960 г. |
| Отношение . . . . . | 621.385.2 : 621.317.7 | Ламповые диоды в электро- и радиоизмерительных приборах |
| Специальный определитель с дефисом . . . . | 621.385.2-181.4 | Ламповые диоды в миниатюрном исполнении |
| Общий определитель точки зрения . . . . . | 621.385.2.001.24 | Расчет ламповых диодов |
| Специальный определитель с точкой и нулем | 621.385.2.019.3 | Надежность ламповых диодов |
| Подразделения простого индекса . . . . . | 621.385.29 | Двойные диоды |

* Индексы со знаком присоединения, например 621.385.2+621.382.2, располагаются как простые индексы. Индексы со знаком объединения (апострофом) располагаются в самом конце, после индексов со специальными определителями.

с той степенью точности и подробности, которая допускается применяемыми таблицами.

*Частная методика* классификации не может быть подробно изложена в рамках настоящей работы. Здесь мы вынуждены ограничиться лишь некоторыми моментами, дающими о ней основное представление. Большая группа вопросов, рассматриваемых частной методикой, касается связи и разграничения между различными отделами схемы.

При классификации документов по естественным наукам и прикладным дисциплинам (разделы 5 и 6 УДК) особенно важно правильно размежевать вопросы между ними. При этом следует учитывать, что естественные науки изучают законы природы, а прикладные дисциплины используют данные естествознания в практических целях. Известно, например, что рентгеновские лучи изучаются в физике, а различные аспекты их применения — в технике и медицине. Поэтому документы, в которых излагаются непосредственные наблюдения или результаты изучения рентгеновских лучей в лабораторных условиях, относятся к отделу 537.531 «Рентгеновские и гамма-лучи. Некорпускулярные лучи при процессах разряда»; документы по конструированию и изготовлению различных рентгеновских аппаратов — к разделу 621.386 «Рентгенотехника», а материалы по диагностическому применению этих лучей — к подразделению медицины 616-073.7. Возможно в данном случае обращение и к другим отделам. Если, например, документ трактует вопрос о фотопленке для рентгеновских снимков, его следует отразить индексом 771.531.34 в разделе фотографии.

Большое значение имеют также вопросы размежевания источников между отделами техники и экономики, а также другими отделами. В частной методике определяется специфика детализации различных отделов, их структура, наиболее целесообразные способы группировок материала по определенным вопросам.

В заключении раздела приводится перечень советских изданий по общей и частной методике классификации по УДК (1963—1966 гг.).

### Пособия по общей методике

Герман П. Практическое применение Универсальной десятичной классификации. Техника классификации. Изд. 5. М., 1967, 97 стр. (ВИНИТИ).

Заполькая Н. Н. Классификация произведений печати и организация систематического каталога в технических библиотеках. М., 1964, 81 стр. (ГПНТБ СССР).

Методика составления алфавитно-предметного указателя к таблицам УДК. М., 1964, 23 стр. (ГПНТБ СССР).

Систематизация справочно-информационных фондов по УДК. Общая методика. М., 1964, 108 стр. (ГПНТБ СССР).

### Пособия по частной методике

Частная методика применения УДК для индексирования *геологических* материалов. М., 1966, 29 стр. (ВИЭМС).

Общая и частная методика индексирования по УДК научно-технической литературы и документации по *горному делу*. М., 1965, 48 стр. (ЦНИИТЭИ угля).

Систематизация справочно-информационных фондов по УДК. *Дорожное дело*. Раздел 629.7/.8. М., 1965, 30 стр. (ГПНТБ СССР).

Частная методика индексации литературы по *кожевенному и меховому производствам* по УДК. М., 1965, 22 стр. (ЦНИИТИ легпром).

Частная методика систематизации литературы по *машиностроению* и смежным вопросам по УДК. М., 1963, 15 стр. (ЦИНТИАМ).

Систематизация справочно-информационных фондов по УДК. *Металлургия*. Раздел 669. М., 1964, 23 стр. (ГПНТБ СССР).

Частная методика индексации литературы по *обувному*, шорно-седельному, перчаточному и кожгалантерейному производствам по УДК. М 1965. 23 стр. (ЦНИИТИ легпром).

Систематизация справочно-информационных фондов по УДК. *Пищевая промышленность*. М., 1966, 19 стр. (ЦИНТИ пищевой промышленности)

Методика классификации *сельскохозяйственной* литературы по УДК М., 1964, 43 стр. (ЦНСХБ ВАСХНИЛ).

Систематизация справочно-информационных фондов по УДК. *Строительство*. Разделы 624 и 69. М., 1964, 15 стр. (ГПНТБ СССР).

Частная методика индексации информационных материалов по *текстильной промышленности* по УДК. М., 1966, 56 стр. (ЦНИИТИ легпром).

Частная методика индексации литературы по *швейному производству* по УДК. М., 1965, 20 стр. (ЦНИИТИ легпром).

## ПРЕДМЕТИЗАЦИЯ

Предметизация — специфический вид систематизации документов, сложившийся в библиотечной практике около столетия назад. Сущность предметизации заключается в том, что содержание документа кратко формулируется при помощи одного или нескольких типовых слов, получивших название предметных рубрик. Затем предметные рубрики располагаются в алфавитном порядке и за ними собираются библиографические описания документов (см. о предметном каталоге в главе 2) или указываются индексы схемы классификации (см. об алфавитно-предметном указателе к систематическому каталогу, там же).

Задачей предметизации является определение того предмета, который в произведении (документе) рассматривается в качестве основной темы, а также установление основного аспекта его характеристики и его отношений и связей с другими предметами. Аспектами характеристики предмета могут служить: его история (происхождение и развитие), структура, состав, свойство, состояние, назначение, взаимодействие с другими предметами, изучение предмета, его оценка и др.

В библиотековедении термин «предметная рубрика» имеет два разных, хотя и связанных между собой значения: 1) краткая словесная формулировка предмета произведения, аспекта его рассмотрения, назначения и формы произведения; 2) структурное деление предметного каталога (указателя), включающее описание произведений, посвященных данному предмету. Было бы более точно говорить о предметной рубрике только во втором значении, применяя для первого термин «предметный заголовок».

В отличие от библиотечно-библиографической классификации, где содержание документа анализируется в отношении его при-

надлежности к определенной отрасли [illegible] с формулировками заранее составленного [illegible] предметизации документы комплектуются [illegible] соотнесенным с какой-либо отраслью [illegible] собрать вместе литературу по такой [illegible] человеческой деятельности, географические [illegible] роды, общественное явление, процесс [illegible] свойство и т. д. Предметизация дает возможность [illegible] документы под такими, например, [illegible] «стекло», «туберкулез», «Париж», «[illegible] сификации документы, содержание [illegible] из этих понятий, будут расположены по [illegible] тественные науки», «Общественные науки», «Технические [illegible] и т. д.).

Другим отличием предметизации от классификации [illegible] то, что, хотя она и производится по заранее составленному списку предметных заголовков, этот список не ограничивает [illegible] анализа содержания документа. Если документ посвящен вопросу, не отраженному в списке предметных заголовков, имеется возможность сформулировать такой заголовок самостоятельно.

Традиционная предметизация, рассчитанная на каталог или картотеку с прямой организацией, или на список, точно так же, как и классификация, не позволяет раскрывать содержание документов во многих аспектах. Предметный заголовок выбирается из списка или формулируется самостоятельно предметизатором лишь по основным аспектам содержания документа. Как при классификации, так и при предметизации один и тот же документ отражается лишь в нескольких рубриках. Рамки предметного комплекса можно расширить посредством ссылок и тем самым достичь некоторого объединения тематически близкой литературы. При достаточно разработанной методике предметизации удается раскрыть наиболее существенные стороны отдельных предметов и их взаимосвязь с окружающей средой.

Самыми важными вопросами методики предметизации являются: применение «широких» и «узких» рубрик, инверсия в их формулировке, использование подрубрик. Не разбирая подробно сущности этой методики, можно на примерах показать решение каждого из вопросов. Предметный заголовок должен возможно точнее определить конкретное содержание документа. Так, например, если в нем трактуется об автомобильных карбюраторах, то рубрика должна формулироваться «Карбюраторы», а не «Двигатели внутреннего сгорания». В этом случае рубрика является «узкой» или, вернее, адекватной, т. е. точно соответствующей основной теме документа. Если более целесообразно документ по определенному вопросу (например, о карбюраторах) отражать под рубрикой, обозначающей более общее понятие (в нашем

примере — «Двигатели внутреннего сгорания»), то такую рубр принято считать «широкой» или, точнее, обобщающей [22].

Инверсия способствует объединению тематически близко материала. Так, например, перестановка прилагательного существительного позволяет сблизить различные виды микр биологии: «Микробиология ветеринарная», «Микробиолог медицинская», «Микробиология промышленная», «Микробиолог сельскохозяйственная». В ряде случаев, наоборот, инверс нецелесообразна, например, когда прилагательное определя существенный признак предмета («Вычислительная техник «Электронные лампы») или в названиях исторических событий других устойчивых словосочетаниях («Парижская коммуна «Детские сады»).

С целью расчленения материала внутри рубрик вводятс подрубрики: тематические («Вода — Анализ, Испарение Очистка, Хлорирование»), географические («Животноводство — СССР, Великобритания, Франция, США»), формальные («Стан дарты — Библиография, Классификаторы, Указатели»), хроно логические («Библиография — История — 19 в.; 20 в.») и др. Особое место занимают стандартные или типовые подрубрики, например, для предметизации литературы о материалах («Бетон — Анализ, Дефекты, Заменители, Испытание»).

Предметизация широко используется в библиотечном и книго издательском деле; на ее основе составляются библиотечные предметные каталоги, алфавитно-предметные указатели к систе матическим каталогам, библиографические указатели, предмет ные указатели к содержанию книг, энциклопедий и справочников с систематическим расположением. Вместе с тем теория предмети зации разработана недостаточно, а методика учитывает лишь отдельные конкретные области ее применения. Нельзя не отметить, что в настоящее время, при создании теории информационно поисковых языков, многие вопросы предметизации, не получив шие научного решения и обоснования, могут быть по-новому освещены в рамках информатики. Такие проблемы координатного индексирования, как установление различных видов связей между дескрипторами, устранение синонимии и омонимии, имеют пря мое отношение к предметизации. Хотя между традиционной пред метизацией и координатным индексированием не было преемст венной связи, их общность не подлежит сомнению.

## АННОТИРОВАНИЕ

Аннотированием называется процесс составления кратких сведений, характеризующих документ со стороны его содержания, идейно-политической направленности, ценности, назначения,

[22] В. П. Кругликова. Предметизация произведений печати. Общая методика. М., «Книга», 1967, стр. 48—49.

оформления и происхождения. Аннотации составляются по тексту документа или на основе других источников с целью углубить информацию о документе, содержащуюся в библиографическом описании.

Аннотации по содержанию и целевому назначению подразделяются на справочные, раскрывающие тематику документа и сообщающие какие-либо сведения о нем, но не дающие его критической оценки, и рекомендательные, содержащие оценку документа в отношении его пригодности для определенной категории потребителей. По охвату содержания аннотируемого документа и читательскому назначению различают аннотации общие, характеризующие документ в целом и рассчитанные на широкий круг потребителей, и специализированные, раскрывающие документ лишь в определенных аспектах, интересующих специалиста более узкой области знаний [23].

В научно-информационной деятельности особо важное значение имеют справочные аннотации. Они могут быть совсем краткими, состоящими из нескольких слов или небольших фраз, и развернутыми, но в отличие от рефератов дают в сжатой форме лишь самые основные положения и выводы документа. Именно эта их черта делает аннотации особенно пригодными для раскрытия во вторичных изданиях (каталогах, библиографических указателях и реферативных журналах) содержания книг. Это объясняется тем, что многоплановое содержание книг с трудом поддается реферированию, тогда как при помощи краткой аннотации можно охарактеризовать не только тематику и структуру книг, но и основные идеи.

В справочной аннотации обычно отражают тип книги и ее назначение (монография, диссертация, справочник, очерк и т. п.), задачи, поставленные автором, и метод, которым он пользовался (эксперимент, сравнительный анализ, компиляция других источников), принадлежность автора к определенной научной школе или направлению, структуру книги, предмет или тему произведения, его основные положения и выводы, характеристику вспомогательных и иллюстративных материалов, дополнений, приложений, справочного аппарата, включая указатели и библиографию. Характерной особенностью аннотации является то, что она должна быть тесно связана со сведениями, включенными в библиографическое описание.

При составлении аннотаций необходимо учитывать определенные требования. Структура аннотации должна быть внутренне логична и может отличаться от структуры аннотируемого произведения. Процесс аннотирования нецелесообразно полностью стандартизировать, как это делается при библиографическом

---

[23] Е. И. Шамурин. Методика составления аннотаций. М., Всесоюз. кн. палата, 1959.

описании. При отборе сведений для аннотации, их формулировании и расположении необходимо учитывать содержание аннотируемого произведения, назначение аннотации и потребителя, на которого она рассчитана. Аннотация должна быть краткой, написанной простым и ясным языком, исключающим длинные и сложные периоды, а также формальные штампы. Средний объем аннотации обычно не превышает 400—600 знаков и лишь в исключительных случаях достигает 800—1000 знаков.

При аннотировании отдельных видов изданий необходимо учитывать некоторые специфические моменты. Для изданий научного характера особенно важно отмечать наличие документальных материалов, статистических данных и справочно-вспомогательного аппарата книги, использовать сведения из вступительных статей и редакционных аннотаций, учитывать отрасль знания, к которой относится содержание книги. В производственной литературе важно подчеркнуть форму произведения (практическое пособие, производственная или методическая инструкция и т. п.), точно соблюдать принятую терминологию и наименования машин и процессов. Аннотируя справочные издания, следует охарактеризовать их тематику, полноту и новизну сообщаемых сведений, источники, на основе которых они составлены, читательское и целевое назначение, принцип расположения в них материала. Аннотация сборника статей или материалов может ограничиваться общим указанием его тематики, сопровождаться перечислением авторов, а также заглавий включенных в сборник статей. Если этого недостаточно, можно дать развернутую характеристику наиболее важных статей, вошедших в сборник, а в случае необходимости и всех статей сборника.

Информация о содержании и характере документа является задачей не только аннотации, но и реферата. Особенности аннотации и реферата связаны с их различным назначением. Аннотация служит только для осведомления о существовании документа определенного содержания и характера, в реферате излагается содержание документа с характеристикой методов исследования или технологии производства, с фактическими данными и итогами работы. В аннотировании основное заключается в умении лаконично обобщить содержание документа, реферирование же предполагает владение мастерством сокращения текста первичного документа.

## РЕФЕРИРОВАНИЕ

Реферирование представляет собой краткое изложение содержания документа с указанием на характер документа, методику исследования и его результаты, а также время и место проведения исследования.

Различают два основных вида рефератов: информативный (реферат-конспект) и индикативный (указательный, или реферат-резюме). Информативный реферат содержит в обобщенном виде все основные положения первичного документа, иллюстрирующего их материал, важнейшую аргументацию, сведения о методике исследования, использовании оборудования и сфере применения. В индикативном реферате приводятся не все, а лишь те основные положения, которые тесно связаны с темой реферируемого документа. Все второстепенное для данной темы в индикативном реферате опускается.

Рефераты, составленные по одному источнику, называются монографическими, а по нескольким источникам на одну тему — обзорными. Среди многочисленных видов рефератов следует выделить специализированные рефераты, в которых изложение ориентировано на специалистов определенной области или какой-либо род деятельности (государственное управление, промышленность и т. п.), а особое внимание уделено одной стороне (оборудованию, технологии, методам исследования). С точки зрения составителя реферата различают авторские рефераты, рефераты, составленные специалистами, по образованию и опыту не уступающими автору оригинала, и, наконец, рефераты (их подавляющее большинство), вышедшие из-под пера профессиональных референтов, эрудированных в широкой области той или иной отрасли знания. Каждый из этих видов имеет свою специфику. При всем разнообразии видов рефераты обладают некоторыми общими чертами.

В реферате, как правило, не должны отражаться субъективные взгляды референта на излагаемый вопрос, не дается обычно и оценка реферируемого материала. В случае крайней необходимости (при явных ошибках или противоречивых положениях автора) такая оценка выносится в примечание.

Если в первичном документе главная мысль сформулирована недостаточно четко, в реферате она должна быть конкретизирована и выделена. Исторические экскурсы в реферате обычно опускаются, а ссылки на опубликованные работы сохраняются лишь в тех случаях, когда это связано с установлением приоритета, когда документ является продолжением ранее опубликованных материалов или в нем эти материалы обсуждаются. При составлении реферата не обязательно придерживаться какого-либо стандартного расположения сведений, но в большинстве рефератов легко прослеживается более или менее единообразный порядок.

Реферат начинается с изложения существа документа, приводится цель работы и сведения о методике и ее сравнительной точности, пересказывается содержание документа, даются выводы автора и указанные им возможности и пути практического применения результатов работы, указывается название и местонахож-

дение учреждения, в котором проводилось исследование. В заключение делаются ссылки на наличие библиографии и иллюстративного материала, если только они не содержатся в библиографическом описании. Цифровые данные, особенно если их много в оригинале, при реферировании систематизируются и обобщаются. Таблицы, схемы, графики, формулы включаются в реферат лишь в той мере, в какой это необходимо для уяснения основного содержания документа. Единицы измерения приняты приводить в соответствии с принятыми в СССР стандартами, географические названия — в русской транскрипции, а иностранные фамилии и названия фирм — в оригинальном написании.

К реферату предъявляется ряд требований. По мнению большинства специалистов, он должен сообщать основные сведения о документе, приводить наиболее ценную информацию из него, давая возможность определять, необходимо ли обращение к первичному документу, быть пригодным к использованию в различных изданиях вторичной информации и механизированных информационно-поисковых системах. Не лишены интереса лаконично сформулированные требования к реферату, публикуемые в каждом выпуске РЖ «Biological Abstracts»:

Реферат должен быть некритическим, информационным резюме существенного содержания и выводов статьи, а не простым ее описанием. Он должен быть понятным сам по себе, без обращения к оригиналу, но не предназначен его заменять. Реферат должен быть краток (желательно не более 3% оригинала) и написан полными предложениями, а не телеграфными фразами.

В качестве примера требований, предъявляемых к реферату в различных реферативных журналах, приводим также выдержки из инструкций для референтов РЖ ВИНИТИ и «Chemical Abstracts».

**Инструкция для референтов и редакторов реферативного журнала ВИНИТИ (М., 1963, стр. 3—4)**

1. Научные материалы освещаются в РЖ в форме рефератов (Р), аннотаций (А) и библиографических описаний (Б). Форма и характер информации (объем реферата или аннотации, степень детализации) должны определяться теоретическим и практическим значением реферируемой работы, ее конкретным содержанием.

2. Основной формой публикации в РЖ является реферат. Реферат представляет собой вторичную публикацию, кратко излагающую содержание первичной публикации. Он должен дать читателю представление о характере освещаемой работы, методике проведения исследования и его результатах, а в случае необходимости (когда такое сообщение имеет научный интерес) — о месте и времени проведения исследований. По реферату читатель должен ясно представить себе степень необходимости обращения к оригиналу. В реферате объективно излагается содержание первичной публикации. Реферат не должен отражать субъективных взглядов референта по излагаемому вопросу, а также оценки референтом освещаемой им публикации. В случае необходимости такая оценка может быть дана в примечании (например, в случае наличия сведений о противоположных результатах в аналогичном исследовании, опыте).

Рефераты составляются на оригинальные работы, представляющие самостоятельный научный интерес, содержащие ранее не опубликованные данные, излагающие новые оригинальные взгляды, новые методы исследования или важнейшие результаты производства. Такими работами являются научные статьи, книги, брошюры и другие непериодические издания; патенты, доклады, прочитанные на конгрессах, конференциях, симпозиумах, дискуссиях; ценные обзорные и некоторые исторические статьи; переводы работ с редких языков на русский, английский, французский, немецкий, если они не реферировались и не переводились ранее; рецензии, если они содержат ценный критический материал или сообщают о публикациях, ранее не известных советскому читателю.

Объем реферата должен соответствовать значению и характеру реферируемой работы. Средний рекомендуемый объем реферата — 1200 печатных знаков. В отдельных случаях (в зависимости от научной ценности реферируемой работы, а также малой доступности издания, где она опубликована) объем реферата может быть увеличен в два-три раза.

3. Аннотация представляет собой краткое изложение содержания публикации без указания экспериментальных данных и методики работы. Рекомендуемый объем аннотации 10—12 строк машинописного текста.

Аннотации составляются преимущественно на справочники, стандарты, путеводители, заметки, обзоры, статьи научно-популярного характера, сообщения о съездах, конференциях, совещаниях (если в них не содержится подробного изложения докладов), рецензии, учебные пособия.

4. Первичные публикации освещаются в РЖ в виде библиографического описания только в крайнем случае. Библиографическое описание представляет собой заглавие работы с точными выходными данными того издания, где она напечатана, и указанием страниц.

### Инструкция для референтов и отраслевых редакторов «Chemical Abstracts»[24]

Характер реферата. 30. Хотя характер и размер рефератов, предназначенных для «Chemical Abstracts», могут быть различными в зависимости от типа реферируемой статьи и определяются в значительной степени самими референтами, рекомендуется принимать во внимание следующие соображения.

31. В реферате следует всегда указывать назначение реферируемой статьи и отражать все содержащиеся в ней новые материалы, представляющие интерес для химиков. Даже в подстрочных примечаниях могут содержаться сведения, заслуживающие упоминания в реферате.

32. Реферат должен быть как можно более кратким, с тем, однако, чтобы эта краткость не затрудняла его понимания. Допустимые размеры реферата ограничены.

33. Следует иметь в виду, что в журнале публикуются безличные резюме или рефераты, а не критические статьи или рецензии. Дело в том, что авторы реферируемых статей не могут на страницах «Chemical Abstracts» отвечать на мнения, высказанные референтами. Тем не менее следует отражать лишь статьи, представляющие ценный вклад в достижения химии, учитывая важность этих статей для современных и будущих читателей журнала.

34. Детальность отражения статьи в реферате может зависеть от источника, в котором она опубликована. Более детального отражения, например, заслуживают статьи, доступные сравнительно немногим американским химикам, так как опубликованы в журналах, не имеющих широкого распространения, или написаны на языке, мало известном в США.

---

[24] Directions for abstractors and section editors of «Chemical Abstracts». Columbus, Ohio State Univ., 1952, p. 6—9.

языке, показало, что лишь 12 из них не пользуются этим приемом. В таком известном РЖ, как «Chemical Abstracts», библиографические описания занимают до 13%, в «Fuel Abstracts» — 24%, а в «Biological Abstracts» — 32%, т. е. составляют треть всех публикаций, освещаемых в журнале [25].

Объем реферата — одна из нерешенных и остро дискутируемых проблем. Практически в различных реферативных изданиях установился средний объем реферата от 100 до 200 слов («Science Abstracts» — 100 слов, «Chemical Abstracts» — 136 слов, РЖ ВИНИТИ — 160 слов, «Biological Abstracts» — 200 слов). Редакции одних реферативных журналов, например «Biological Abstracts», ставят объем реферата в некоторую зависимость от величины реферируемой первичной публикации. В других реферативных журналах, например РЖ ВИНИТИ, объем реферата определяется конкретным содержанием первичной публикации [26]. Однако a priori ясно, что на объем реферата влияют и многие другие факторы, например значение и характер реферируемой работы, ее доступность для читателей реферативного журнала. Оригинальные экспериментальные исследования по актуальным проблемам с большой продолжительностью опытов и наблюдений заслуживают более детального реферирования, чем научные обзоры, например научно-популярные или научно-производственные работы.

Однако трудно сказать, насколько оптимальными являются эти эмпирически установленные критерии. Дело в том, что эффективность, действенность рефератов изучена очень плохо, а предъявляемые к ним различные требования часто противоречивы. Среди работников информационных служб нет единства даже по вопросу о том, какая существует зависимость между степенью сокращения первичного документа и потерей существенной информации в реферате. С одной стороны, исследования, проводившиеся в Технологическом институте Кейса (США, 1960 г.) над текстами документов, сокращенными до 2/3, 1/3 величины оригинала и до размеров нормального реферата, показали, что процент потери информации мало существен. С другой стороны, наблюдения сотрудников фирмы International Busines Machines (США, 1961 г.) говорят о том, что свертывание документа в реферат ведет к потере информации от 50 до 75% [27].

Особый подход к проблеме реферирования возникает при попытках автоматизировать этот процесс. Автоматизация рефе-

[25] J. B. Reed. The diversity of text-length amongst systems of scientific and technological abstracts. «Journal of Documentation», 1966, v. 22, № 2, стр. 136—143.

[26] Инструкция для референтов и редакторов реферативного журнала ВИНИТИ. М., 1963, стр. 3.

[27] B. H. Weil, I. Zarember, H. Owen. Technical-abstracting fundamentals. «Journal of Chemical Documentation», 1963, v. 3, № 3, стр. 125—132.

ков, изобилующих ключевыми словами, отобранными по специальным правилам из самого реферируемого текста. Установление новизны излагаемого материала и его критическая оценка в данных науках пока не поддаются автоматизации. Теоретическая разработка проблемы автоматического реферирования в изложенном выше аспекте ведется группой сотрудников ВИНИТИ [28]. Следует отметить, что для решения этой проблемы, наряду с отмеченным подходом, необходимо изучение методики составления рефератов людьми.

Иногда под автоматическим реферированием понимают машинную переработку научно-технических текстов, при которой из них выделяются отдельные предложения, отражающие цели и результаты реферируемой работы. Поскольку набор предложений, получающийся в результате такой переработки, строго говоря, не является рефератом, в данном случае уместнее говорить об «автоматическом квазиреферировании». Одним из первых исследователей, предложивших воспользоваться при рассмотрении этой проблемы понятием «множества значимых слов», был Г. П. Лун. Он писал, что «правомерность измерения значимости слов частотностью их употребления объясняется тем, что каждый автор в процессе рассуждения обычно повторяет определенные слова. Это средство подчеркивания той или иной мысли принимается за указатель значимости. Чем чаще некоторые слова встречаются в сочетании друг с другом в предложениях, тем бóльшую значимость можно приписать каждому из этих слов» [29].

Автоматическое квазиреферирование на основе статистического анализа текста может проводиться по различной методике. В зависимости от этого и результаты такого квазиреферирования будут различными. Г. П. Лун, например, предложил для оценки степени значимости предложений сосчитать число значимых слов в предложении, возвести эту величину в квадрат и затем разделить на общее число слов в предложении. Советские специалисты В. А. Аграев и В. В. Бородин (Горьковский государственный университет) предложили методику автоматического квазиреферирования, при которой выбранные из текста предложения оказываются связанными друг с другом. При этом наиболее связанными они считают те предложения, в которых имеется наибольшее число одинаковых значимых слов.

---

[28] А. М. Васильев. Автоматическое реферирование и индексирование. В кн.: Материалы заседания рабочей группы № 2 ЮНЕСКО по вопросу «Автоматизация информации — поиск и хранение» (Москва, 11—16 ноября 1963 г.). М., ВИНИТИ, 1965, стр. 83.

[29] H. P. Luhn. The automatic creation of literature abstracts (auto-abstracts). «IBM Journal of Research and Development», 1958, v. 2, № 2, стр. 159—165.

В. А. Пурто предложил для этой цели метод оценки и отбора предложений по количеству содержащейся в них информации [30]. По его методу тексты подвергаются статистическому анализу для выяснения частотности употребления входящих в них слов. Наиболее часто употребляемыми словами в научно-технических текстах (служебные слова во внимание не принимаются) оказываются термины. В. А. Пурто исходит из априорного предположения, что чем важнее термин, тем чаще он встречается в тексте. Поэтому он считает, что предложения, отобранные по его методу, будут содержать много важных терминов. Полученный таким образом квазиреферат несколько напоминает аннотацию, объем которой не зависит от объема реферированного текста и не превышает двух-трех предложений.

Для примера приведем реферат статьи, опубликованной в РЖ ВИНИТИ «Машиностроение», и квазиреферат той же статьи, автоматически составленный по методу В. А. Пурто.

**Реферат статьи И. Н. Величкина**
**«Особенности рабочего процесса газогенераторного двигателя Г-58»**

«Вестник машиностроения», 1956, № 8, стр. 6—11,
выполненный специалистом-референтом и опубликованный
в РЖ «Машиностроение», 1957, № 18, реф. № 38477

Высокое объемное содержание горючего в смеси генераторного газа с воздухом является причиной существенной разницы между зависимостями отдельных параметров рабочего процесса двигателя от коэффициента избытка воздуха ($\alpha$) при питании генераторным газом и жидким топливом нефтяного происхождения. Для газогенераторного питания характерно (по сравнению с питанием жидким топливом): получение максимума индикаторного к. п. д. при значительно большем коэффициенте избытка воздуха; возможность значительного увеличения коэффициента избытка воздуха при достаточно устойчивой работе двигателя; повышение коэффициента наполнения по мере увеличения коэффициента избытка воздуха. Экспериментальная проверка выполнена испытанием газового тракторного двигателя Г-58 при питании его генераторным газом древесного угля с присадкой воды. Приведены графики, характеризующие влияние коэффициента избытка воздуха на параметры рабочего процесса. Расширение пределов горючести генераторного газа за счет обеднения смеси служит предпосылкой к использованию качественного регулирования хотя бы на части поля возможных изменений нагрузки. Сопоставление нагрузочных характеристик при качественном и количественном регулировании подтвердило экономическое превосходство первого в зоне высоких нагрузок.

---

[30] В. А. Пурто. Об автоматическом реферировании на основе статистического анализа текста. «Докл. на Конференции по обработке информации, машинному переводу и автоматическому чтению текста», вып. 9. М., 1961, стр. 10 (АН СССР, Ин-т научной информации).

## Квазиреферат статьи И. Н. Величкина «Особенности рабочего процесса газогенераторного двигателя Г-58»

(выполнен по методу В. А. Пурто)

В связи с этим были поставлены опыты по выяснению следующих вопросов: 1) влияние малого объемного коэффициента смешения на параметры рабочего процесса газогенераторных двигателей; 2) возможности использования широких пределов горючести генераторного газа для совершенствования конструкции двигателей; 3) возможности увеличения мощности газогенераторных двигателей путем уменьшения значения $\alpha$, соответствующего максимальной мощности.

Более высокие значения индикаторного к. п. д. при качественном регулировании по сравнению со значениями при количественном регулировании являются следствием:

а) увеличения коэффициента избытка воздуха в смеси, что ведет к уменьшению многоатомных газов в продуктах сгорания и, следовательно, к уменьшению их теплоемкости;

б) уменьшения температуры цикла, что также снижает теплоемкость отработанных газов;

в) уменьшения коэффициента остаточных газов вследствие более полного пополнения цилиндров двигателей свежей горючей смесью при качественном регулировании; уменьшение коэффициента остаточных газов обеспечивает более быстрое сгорание смеси, на что указывает меньшая температура отработанных газов.

Таблица 12

**Результаты автоматического квазиреферирования в зависимости от методики**

| Метод автоматического квазиреферирования | Порядковые номера отобранных предложений |
|---|---|
| Оценка степени значимости предложений по отношению квадрата числа значимых слов предложения к общему числу слов в предложении (метод Г. П. Луна) . . . . . | 16, 19, 54, 57, 59, 64 |
| Оценка степени значимости предложений по наибольшему числу встречающихся в них одинаковых значимых слов (метод В. А. Аграева и В. В. Бородина) | |
| с учетом частоты отдельных значимых слов | 10, 16, 19, 61, 83, 84, 89 |
| с учетом частоты употребления последовательных словосочетаний значимых слов . . . . . | 10, 16, 19, 34, 54, 101, 103 |
| с учетом частоты употребления всех возможных внутрифразовых сочетаний значимых слов | 10, 15, 16, 19, 59, 61, 101 |
| Оценка значимости предложений по частоте употребления входящих в них значимых слов (метод В. А. Пурто) . . . . . | 19, 101 |

Различие результатов автоматического квазиреферирования цитированной статьи И. Н. Величкина в зависимости от применяемой методики можно проиллюстрировать табл. 12.

Необходимо особо остановиться на методике обзорного реферирования. Известно, что реферативный обзор отличается от аналитического меньшей глубиной анализа документов. В нем, как правило, даются только систематизация и обобщение сведений из первичных источников. Реферативный обзор может не содержать аргументированной оценки материала и обоснованных выводов для его практического использования. Тем не менее правильно составленный реферативный обзор является ценным видом информации, позволяющим сосредоточить внимание на наиболее важных документах.

При составлении реферативного обзора уясняются тематические и временные границы используемого материала, определяются значение и актуальность темы, содержание и методы теоретических разработок, сведения технологического, экономического и организационного характера. В процессе обзорного реферирования используются документы и материалы всех видов, если они содержат новейшие сведения, факты и идеи: постановления и распоряжения правительственных учреждений, отечественные и зарубежные периодические издания, книги, отчеты о научно-исследовательских и опытно-конструкторских разработках, частные технические решения из неопубликованных документов, описания отечественных и зарубежных изобретений, ведомственные, фирменные и рекламные издания, диссертации, материалы выставок, отчеты о командировках, технические условия и нормали, материалы совещаний, симпозиумов, конгрессов и т. п.

Во введении к обзору рассматривается значение и краткая история вопроса, указывается взаимосвязь со смежными областями, определяется читательское назначение обзора. Текст обзора представляет собой последовательное, логически связанное изложение идей и фактов. Он должен быть не механическим пересказом источников, а творческим синтезом, обобщением содержащихся в них важнейших сведений. В случае необходимости обзор иллюстрируется фотографиями, графиками, диаграммами, чертежами, функциональными, принципиальными или монтажными схемами. В заключение подводится сравнительный итог главных положений и сведений обзора, но без их критической оценки и конкретных выводов. Выявляются лишь общий уровень и тенденции развития данной области или вопроса. Обзор заканчивается списком использованных источников, расположенных в той последовательности, в которой они упоминались в нем. Обзоры объемом свыше одного авторского листа целесообразно сопровождать оглавлением.

## НАУЧНЫЙ ПЕРЕВОД

Перевод различных текстов с одного языка на другой возник из потребностей производственной и общественной деятельности человека в связи с ролью языка как средства общения. История переводческой деятельности восходит к глубокой древности. В античные времена довольно часто переводились на греческий и латинский языки различные политические документы, религиозные тексты и художественные произведения. Однако изучать явления перевода стали сравнительно недавно, причем основное внимание уделялось теории и методике художественного перевода.

Перевод текстов, не относящихся к художественной литературе — научных, технических, политических, экономических и т. п., — мы условно называем научным переводом. В связи с бурным ростом научных публикаций вопросы научного перевода приобрели особенно большое значение. Если в течение нескольких веков латинский язык выполнял роль международного языка науки, а в прошлом столетии ученому было вполне достаточно знания английского, немецкого и французского языков, то в наши дни без учета научных публикаций на русском и ряде других языков успешная научная работа невозможна. Это влечет за собой необходимость переводов научных текстов с одного языка на другой.

По данным международного указателя переведенных книг «Индекс Транслационум», 10% всех поступающих на мировой книжный рынок названий составляют переводы. В 1965 г. было зарегистрировано 36 588 названий переводных книг, причем в СССР опубликовано наибольшее их число — 3864, т. е. около 10% общего числа переводов [31]. Значительная доля переведенных книг падает и на произведения советских авторов, издающиеся за рубежом. В США, например, на переводы с русского языка по естественным наукам приходится 45%, а по прикладным наукам — 40% всех переведенных книг. Около ста названий журналов полностью и регулярно переводится с русского на английский язык, публикуется и распространяется по подписке в США, Великобритании и других странах. Десятки названий журналов переводятся с иностранных языков на русский и издаются в нашей стране. Десятки тысяч переводов выполняются в научных учреждениях и на промышленных предприятиях; как правило, они не публикуются и почти никак не учитываются. Если же принять во внимание трудности, связанные с отсутствием квалифицированных переводчиков-специалистов, несовершенством специальных словарей, неунифицированностью терминологии, неразработанностью проблем теории и практики научного перевода, то можно понять, какая огромная масса духовных ресурсов чело-

[31] Index Translationum, v. 18. Paris, Unesco, 1967.

вещества растрачивается нерационально в безуспешных попытках преодолеть языковый барьер в области научных публикаций.

«Современный мир выглядит как огромная переводческая машина, работающая со все увеличивающейся скоростью. Каждый день с печатных станков сходит около 60 переведенных книг, открывается 3 или 4 международных встречи, выпускается на экран несколько дублированных фильмов; пишущие машинки печатают, а дикторы оглашают несчетное количество переведенных страниц; бесконечное число документов переводится в самом разнообразном виде с одних языков на другие» [32].

Вместе с тем публикуемые в настоящее время переводы далеко не удовлетворяют потребности ученых и специалистов различных стран во взаимном ознакомлении с достижениями науки и техники.

В связи с возрастающими трудностями языкового барьера в передаче научной информации все большее распространение стали получать исследования по автоматическому (машинному) переводу. Впервые мысль о возможности такого перевода была высказана еще три столетия назад Г. Лейбницем. В 1933 г. советский изобретатель П. П. Троянский получил патент на переводческую машину. Однако датой рождения идеи автоматического перевода в ее современном понимании считается 1946 г., когда английский ученый А. Д. Бут в дискуссии с американцем У. Уивером высказал мысль о возможности использования любой электронной цифровой вычислительной машины для пословного перевода [33]. В 1949 г. У. Уивер опубликовал свой знаменитый меморандум «Перевод», первую работу о возможности перевода с одного языка на другой при помощи ЭЦВМ [34]. С этого времени проблема автоматического перевода начинает широко обсуждаться и разрабатываться американскими и английскими учеными. В 1954 г. в США был осуществлен первый публичный опыт автоматического перевода. В 1955—1956 гг. подобные опыты были проведены в нашей стране. С тех пор над разрешением этой проблемы упорно работают лингвисты, математики, специалисты в области вычислительной техники.

Первоначально основным содержанием этих исследований считалось построение алгоритмов для перевода с одного конкретного языка на другой при помощи электронной машины определенного типа. Большинство исследователей тогда составляли программисты, математики, логики, инженеры. С конца 50-х годов стало складываться более широкое понимание автомати-

[32] E. Cary. La traduction dans le monde moderne. Genève, 1956, стр. 62.

[33] A. D. Booth. A progress report on machine translation. «Bull. du Centre international provisoire de calcul (Rome)», 1961—1962, № 15—16, стр. 11—19.

[34] W. Wever. Translation. В кн.: «Machine translation of languages». New York—London, 1955, стр. 15—23.

ческого перевода (АП) как проблемы в основном лингвистической, хотя теории и методы традиционного языкознания оказались недостаточно разработанными для ее решения.

«Вместе с тем АП начинает трактоваться как частная задача в рамках более общей и научно более привлекательной задачи — обучения электронных машин человеческим языкам, т. е. моделирования речевой деятельности в целом. При этом АП в буквальном смысле слова никак не может считаться первым этапом на пути решения этой общей задачи: перевод с одного языка (входного) на другой (выходной) не является простейшим видом речевой деятельности. Умение переводить обязательно предполагает: 1) умение понять текст на входном языке, т. е. извлечь из него смысл и определенным образом зафиксировать этот смысл; и 2) умение выразить данный смысл на выходном языке. Сложный процесс перевода распадается на более простые процессы „понимания" (анализ) и „выражения" (синтез). Следовательно, моделирование речевой деятельности целесообразно начинать не с перевода, а с его обязательных компонентов, т. е. с „понимания" и „выражения". . . Серьезное изучение, моделирование и описание владения языком — т. е. речевой деятельности „в лице" понимания и выражения — вполне соизмеримы по сложности и значимости с такими проблемами современной науки, как расшифровка генетического кода или теория элементарных частиц» [35].

Хотя до сих пор машинный перевод еще не может быть практически использован, исследования в этой области сыграли значительную роль в развитии лингвистики, в первую очередь теории перевода.

«Машинный перевод — это первый шаг к объективному эксперименту в теории перевода и вообще в языкознании. Машинный перевод предполагает те или иные гипотезы о процессе перевода, которые в дальнейшем проверяются путем моделирования процесса перевода в машине. Ясно, что при такой постановке вопроса машинный перевод играет не только практическую роль, он открывает пути для объективного изучения языковых процессов, имеющих место в переводе. Это, в свою очередь, должно обогатить и практическую реализацию машинного перевода, а также помочь построению других машин, связанных с языком: машин для перевода устной речи в письменную, информационных машин и т. п.» [36].

Совершенно естественно, что проблемы общего и машинного перевода далеко выходят за рамки вопросов информатики. Однако

---

[35] И. А. Мельчук. Предисловие к кн.: И. А. Мельчук, Р. Д. Равич. Автоматический перевод. 1949—1963. Критико-библиографический справочник. М., 1967, стр. 7 (Всесоюз. ин-т научной и техн. информации. Ин-т языкознания).

[36] И. И. Ревзин, В. Ю. Розенцвейг. Основы общего и машинного перевода. М., «Высшая школа», 1964, стр. 18.

они имеют несколько точек соприкосновения с проблемами научной информации. Прежде всего в теоретическом плане успехи машинного перевода открывают новые возможности для создания информационно-логических систем, способных извлекать новые научные сведения из уже известной информации. Очевидно, что после успешной практической реализации машинного перевода соответствующие устройства станут частью информационно-поисковых и информационно-логических систем.

В настоящее время процесс перевода научных и специальных текстов с одного языка на другой является неотъемлемым этапом информационной деятельности любого органа научной информации. В связи с этим для десятков тысяч информационных работников, являющихся профессиональными переводчиками, исключительно важно учитывать в своей работе достижения теории и методики общего и специального перевода с одного языка на другой. Существуют такие области прикладной лингвистики, которые, наряду с языковедами, успешно разрабатываются специалистами-практиками, в том числе и работниками информационных служб.

Крупные центры информации получают значительное число публикаций на десятках различных языков. Для того чтобы оперативно направлять эту литературу и другие информационные документы по многочисленным каналам аналитико-синтетической переработки, необходимо на начальном этапе иметь возможность быстро и точно определять, на каком языке написан текст. Понятно, что сосредоточение на этом этапе специалистов по многим языкам нерентабельно и не всегда осуществимо. Для реализации этой задачи прибегают к специальным пособиям, справочникам по языкам или, как их называют, определителям языков. Возможность определения незнакомого иностранного языка по внешним признакам основывается на том, что почти каждый язык имеет в своем письме специфические, одному ему присущие графические особенности, отличающие его от всех других или от сходных с ним языков. Многие письменности используются только в одном языке, наиболее распространенные — в ряде языков, которые в пределах этой письменности имеют различные алфавиты, особые буквы, значки при буквах или служебные слова, не встречающиеся в других языках. Пользуется известностью несколько пособий подобного рода. Одним из последних по времени издания является вышедший в нашей стране третьим изданием на русском [37] и английском [38] языках «Определитель языков мира по письменностям». Он охватывает свыше 200 языков

[37] Р. С. Гиляревский, В. С. Гривнин. Определитель языков мира по письменностям. Изд. 3, испр. и доп. М., «Наука», 1965.

[38] R. S. Gilyarevsky, V. S. Grivnin. Langauages identification guide. M., «Nauka», 1968.

Европы, Азии, Океании, Африки и Америки, на которых издается подавляющее большинство выходящей в мире литературы.

В процессе научного перевода с одного языка на другой, реферирования и аннотирования иноязычных текстов часто встречается необходимость передачи иностранных имен собственных средствами русского алфавита. Основным способом такой передачи служит *практическая транскрипция*. Под транскрипцией понимают передачу звуков или начертаний языка системой знаков, отличных от принятых в этом языке письменных единиц. Эта система знаков может быть условной, например при фонетических транскрипциях, но может представлять собой и исторически сложившуюся орфографическую систему другого языка. Последний случай принято называть практической транскрипцией.

Транскрипции подлежат различные собственные имена: фамилии и имена лиц, географические наименования, названия небесных тел, учреждений, организаций, фирм, журналов и газет, морских и речных судов, самолетов и т. п., а также многие специальные термины. В 1931 г. академик Л. В. Щерба назвал состояние русской транскрипции «хотя и маленьким, но все же общественным бедствием» [39]. С тех пор у нас была проведена значительная работа по унификации транскрипции, прежде всего географических названий. Главное управление геодезии и картографии, начиная с 1940 г., выпускает инструкции по передаче на картах географических названий различных стран, способствующие унификации этого вида собственных имен.

Трудность практической транскрипции заключается в том, что она производится по-разному не только в зависимости от языка, звуки которого передаются русскими буквами, но и в зависимости от вида имен собственных. Некоторые виды имен собственных, например астрономы (названия небесных тел), всегда переводятся, другие, например названия государств, большей частью передаются в традиционном написании, третьи, например названия организаций, частично транскрибируются, частично переводятся. Транскрипция многих известных названий и имен осуществляется в традиционном написании, отражающем старые транскрипционные нормы. Для того чтобы в процессе информационной деятельности однозначно передавать на русском языке иностранные названия и личные имена, приходится вести соответствующие картотеки. Существенную помощь в этой работе могут оказать практические пособия по транскрипции. Последнее такое пособие издано в 1968 г.[40]

---

[39] Л. В. Щерба. Транскрипция иностранных слов и собственных имен и фамилий. «Изв. Комиссии по русскому языку АН СССР», 1931, т. 1, стр. 187.

[40] Р. С. Гиляревский и Б. А. Старостин. Иностранные имена и названия в русском тексте. Справочник. М., «Международные отношения», 1968.

При библиографическом описании, составлении библиографических указателей, организации каталогов и картотек и в других процессах информационной деятельности возникает необходимость передачи русских имен собственных, главным образом фамилий и личных имен, средствами письменностей иностранных языков. Для подобной передачи используют частный случай транскрипции — *транслитерацию*. «В отличие от транскрипции, которая передает звуки одного языка буквами или сочетаниями букв другого языка, транслитерация может быть определена как точная передача знаков одной письменности знаками другой письменности» [41].

Проблема написания русских фамилий и названий латинскими буквами разрабатывалась русской лингвистической наукой более 50 лет. В ее решении большая роль также принадлежит академику Л. В. Щербе, работы которого легли в основу транслитерационной системы АН СССР, созданной под его непосредственным руководством [42]. Эта система была почти полностью повторена в соответствующей рекомендации Международной организации по стандартизации (МОС—$Р_9$—1954), которая в настоящий момент переработана. Недостатком этой системы являлось то, что она охватывала лишь шесть языков, пользующихся русским письмом. Между тем этим письмом пользуется свыше 60 языков народов СССР, на которых издается основная масса выходящей в нашей стране литературы. Для транслитерации латинскими буквами текстов на языках народов СССР в 1960 г. была предложена система, разработанная во Всесоюзной книжной палате [43]. Эта работа положена в основу проекта пересмотра соответствующей рекомендации МОС [44]. Она в настоящее время выступает в качестве единственной системы, практически применяемой в международных изданиях (например, в «Index Translationum») для транслитерации собственных имен авторов, пишущих на языках народов СССР, пользующихся русским письмом. На заседании Технического комитета МОС «Конверсия письменных языков» во время Генеральной ассамблеи МОС в Москве в июне 1967 г. окончательный вариант этой системы обсуждался в качестве рабочего документа.

---

[41] Н. В. Юшманов. Ключ к латинским письменностям земного шара, М.—Л., 1941, стр. 19.

[42] Л. В. Щерба. Транслитерация латинскими буквами русских фамилий и географических названий. «Изв. АН СССР. ОЛЯ», 1940, № 3, стр. 118—126.

[43] Р. С. Гиляревский и Н. В. Крылова. Транслитерация библиографических описаний на языках народов СССР латинскими буквами. «Сов. библиография», 1960, вып. 5, стр. 37—44.

[44] Международная организация по стандартизации. Технический комитет 46 — Документация. Отчет о заседании в Брюсселе 14—16 октября 1963 г. Резолюция 2.

Новый вариант транслитерации букв русского алфавита, принятый МОС в 1967 г., мало отличается от предыдущего. Он приведен в табл. 13.

Таблица 13

**Транслитерация букв русского алфавита**

| | | | |
|---|---|---|---|
| а—a | и—i | р—r | ш—š, sh |
| б—b | й—j | с—s | щ—šč, shch |
| в—v | к—k | т—t | ъ—″ |
| г—g | л—l | у—u | ы—y |
| д—d | м—m | ф—f | ь—′ |
| е—e | н—n | х—h, kh, ch | э—ė |
| ё—ë | о—o | ц—c, ts | ю—ju, yu |
| ж—ž, zh | п—p | ч—č, ch | я—ja, ya |
| з—z | | | |

Особенностями нового варианта является наличие альтернативных вариантов передачи букв ж, х, ц, ч, ш, щ, ю, я. Из них первые шесть могут передаваться либо отдельными буквами (некоторые с диакритическими знаками), либо диграфами. При этом для отличения диграфов от сочетания соответствующих букв применяется точка, например: ц—ts, тс—t · s.

Для транслитерации текстов на языках неславянских народов СССР, пользующихся русским письмом, применяются следующие общие принципы: 1) буквы, совпадающие с буквами русского алфавита, передаются аналогично с ними; 2) буквы, которых нет в русском алфавите, задаются особой таблицей (табл. 14).

Необходимо отметить, что за последние три десятилетия практические нужды транслитерации русских текстов вызвали к жизни несколько транслитерационных систем, из которых основными являются приведенная выше рекомендация МОС (с диакритическими знаками) и англо-американский стандарт (без диакритических знаков, с диграфами). Однако до последнего времени проблема транслитерации не подвергалась дальнейшей теоретической разработке. В 1965 г. американский лингвист Р. О. Якобсон выдвинул принцип автоматической передачи букв одного алфавита буквами другого алфавита и соответствия транслитерационных систем требованиям точности, единообразия и экономии в телеграфной связи [45].

[45] Р. О. Якобсон. О латинизации международных телеграмм на русском языке. «Вопросы языкознания», 1965, № 1, стр. 111—113.

**Транслитерация букв алфавитов языков народов СССР**

| | | | | |
|---|---|---|---|---|
| æ, ə — ä | ӂ, ӝ — dž | қ, к, ҡ, ҟ, ҳ, ҷ — q | ң — ń | ӳ, ӯ, ӱ, ү — ü |
| á—aa | ц — c | | й — ŋ | ҫ — ç |
| ɜ—â | ч — č | | о — ö | ҵ — dč |
| ґ, г, ғ — ğ | ҙ — dz | | ó — oo | ҽ — tš |
| h, ӏ — ’ | ҙ — dź | | ҧ — ṕ | ӵ — č |
| ҳ—h | з — d’ź | | ҫ — ş | ’ — ‘ |
| ҽ—čč | л’—l’ | | ҫ — ś | |
| ҿ—dčdč | ӆ — l̨ | | ҭ, ҭ, ҭ — t’ | |
| | й—ii | ң, ҥ, ң, ӈ — ng | | |
| | ӣ—í | | | |
| | j—j | | | |

Знак (¯) над гласной означает ее удвоение, например, ā—aa
Знаки (˘), (ˇ) над гласной сохраняются, например, ă—ă, ǎ—ǎ

Новый шаг в научном изучении принципов транслитерации был сделан советским семиотиком В. А. Успенским в 1967 г.[46] Прежде всего он по-новому подошел к определению лингвистических понятий «перевод», «практическая транскрипция» и «транслитерация». Он считает, что эти термины обозначают различные способы передачи слов, словосочетаний и вообще текстов («цепочек») одного языка («языка-источника») средствами другого языка («языка-восприемника»). Переводом он называет передачу *значения* цепочки, когда в языке-восприемнике подбирается такая цепочка, одно из значений которой было бы, насколько это возможно, похожим на одно из значений исходной цепочки языка-источника.

Практической транскрипцией В. А. Успенский считает передачу *звучания* цепочки, когда в языке-восприемнике подбирается такая цепочка, одно из звучаний которой было бы, насколько это возможно, похожим на одно из звучаний исходной цепочки в языке-источнике. «Таким образом, если даны две цепочки — одна в языке-источнике, другая в языке-восприемнике, то можно обсуждать вопрос, является ли вторая переводом или практической транскрипцией первой; вопрос, является ли вторая

[46] В. А. Успенский. К проблеме транслитерации русских текстов латинскими буквами. «Научно-техн. информация», 1967, серия 2, № 7, стр. 12—20.

цепочка транслитерацией первой, остается беспредметным до тех пор, пока не указана система транслитерации»[47].

Транслитерацией в цитированной работе считается передача написания цепочки, когда в языке-восприемнике подбирается такая цепочка, которая в силу специального соглашения (называемого системой транслитерации) считалась бы отражающей написание исходной цепочки в языке-источнике. Поэтому рассматривать транслитерацию той или иной цепочки можно не вообще, а лишь относительно какой-либо системы. В работе впервые строго сформулированы требования, предъявляемые к системе транслитерации. Главными требованиями В. А. Успенский считает: 1) *общеприменимость* — система должна давать латинскую транслитерацию для любой русской цепочки, а не только для значимых слов и текстов; 2) *однозначность* — латинская транслитерация русской цепочки должна быть единственной; 3) *обратимость* — любая русская цепочка должна однозначно восстанавливаться по своей латинской транслитерации; 4) *фонетическую непротиворечивость* — при попытке сделать чтение транслитерированных букв и буквосочетаний похожим на чтение исходных русских цепочек, их фонетическая интерпретация не должна противоречить принятой в наиболее распространенных языках, использующих латиницу. Дополнительные требования, предъявляемые им к системе транслитерации, подобно четвертому главному требованию, не имеют формального характера: 1) *простота системы* — процесс транслитерации не должен занимать много времени; 2) *простота ретранслитерации*; 3) *графическая простота* — система не должна иметь знаков, сложных для печати или легко смешиваемых друг с другом; 4) *экономность* — подвергаясь транслитерации, слово должно как можно меньше растягиваться; 5) *традиционность* — система не должна без необходимости отступать от сложившихся традиций.

Легко заметить, что отдельные требования, сформулированные В. А. Успенским, противоречат друг другу. Поэтому, как справедливо отмечает сам автор, невозможно построить систему, которая удовлетворяла бы каждому из этих требований. Любая система транслитерации строится на компромиссе между ними. Этим объясняется большое число существующих систем и наличие определенных недостатков в каждой из них. Не лишена их и система, предлагаемая в рассматриваемой работе. Однако существенное значение этой статьи для дальнейшей разработки проблемы транслитерации заключается в том, что в ней впервые сформулированы принципы, которые до сих пор лишь интуитивно и далеко не полностью учитывались при составлении конкретных систем транслитерации.

---

[47] Там же, стр. 14.

Глава 4

# ИНФОРМАЦИОННО-ПОИСКОВЫЕ СИСТЕМЫ

## ИНФОРМАЦИОННЫЙ ПОИСК. ОБЩИЕ ПРИНЦИПЫ И ПОНЯТИЯ

Важными внутренними закономерностями развития науки, как уже отмечалось ранее, являются ее преемственность и интернациональный характер. Ни одно крупное научное открытие — независимо от эпохи, в которой оно было сделано, — не является продуктом труда одного человека или даже одного поколения. Оно представляет собой результат деятельности многих людей и многих поколений. Эту мысль в свое время очень образно выразил Исаак Ньютон, написавший: «Если я видел несколько дальше других, то потому, что я стоял на плечах гигантов» [1]. Вероятно, он имел в виду прежде всего Н. Коперника, Тихо де Браге, Г. Галилея и И. Кеплера. Более четко эту мысль сформулировал К. Маркс. «Всеобщим трудом, — писал он в „Капитале", — является всякий научный труд, всякое открытие, всякое изобретение. Он обусловливается частью кооперацией современников, частью использованием труда предшественников» [2]

Преемственность в развитии науки, ее коллективный характер и интернационализм обеспечиваются системой письменных документов, в которых излагаются результаты исследований и разработок как отдельных лиц, так и соответствующих коллективов. Без письменности, письменных документов было бы невозможно возникновение и развитие науки, так как до настоящего времени письменный документ — журнальная статья, научно-технический отчет, патентное описание, монография и т. д. — продолжает оставаться важнейшей формой материального закрепления научной информации с целью ее передачи во времени и пространстве. В предисловии к книге Д. Диринтера «Алфавит» И. М. Дьяконов пишет: «Письменность — одно из самых существенных завоеваний человека. . . Без нее не было бы возможно

[1] Цит. по: Т. Саймон. Поиски планеты Икс. М., «Мир», 1966, стр. 19.
[2] К. Маркс. Капитал, т. 3, М., Госполитиздат, 1955, стр. 109.

то огромное культурное развитие человечества от каменного до атомного века, которое сделало современный мир таким, каким он является ныне. Вплоть до изобретения радио и звукозаписи только письменность давала возможность передавать творения человеческой мысли на расстоянии и сохранять их в течение веков. Однако и сейчас письменность имеет ряд преимуществ перед механической звукозаписью, прежде всего в компактности и обозримости письменного текста»[3]. И далее: «Изобретение письма сделано возможным массовое накопление знаний и передачу любых сообщений, а также произведений художественного слова и науки вне личного общения и на протяжении неограниченного времени»[4].

Наука, как и другие общественные явления, имеет несколько разных сторон, которые неразрывно связаны друг с другом. Прежде всего, наука — это система упорядоченных знаний, форма общественного сознания. Но в то же время наука есть разновидность общественного разделения труда, специфичный вид деятельности людей, направленный на получение новой научной информации на основании той, которой уже располагает человеческое общество. Под научной информацией в данном случае понимается получаемая в процессе познания логическая информация, которая адекватно отображает закономерности объективного мира и используется в общественно-исторической практике[5]).

Основным источником научной информации, необходимой для проведения любого исследования, служат письменные документы; в письменных же документах излагаются результаты этого исследования и дается их общая интерпретация. Для некоторых наук (например, для истории) письменные документы играют еще более важную роль. Во-первых, для историков содержание различных научных документов является одним из основных объектов исследования; во-вторых, изложение результатов исследования в новом письменном документе — статье или монографии — есть основная форма проведения заключительного этапа самого исследования.

На рис. 24 приведена упрощенная схема передачи научной информации в процессе развития науки. Эта схема имеет два существенных недостатка. В ней «информация β» поступает к «исследователю А», как бы замыкая цикл с двумя каскадами обогащения научной информации. Фактически этот цикл представляет собой не замкнутый круг, а виток восходящей спирали. Кроме того, на рис. 24 надписи «исследователь А» и «исследователь Б» нужно заменить надписью «исследователи», имея в виду всех исследователей

[3] И. М. Дьяконов. О письменности. В кн.: Д. Дирингер. Алфавит. М., ИЛ, 1963, стр. 5.
[4] Там же, стр. 6.
[5] А. И. Михайлов, А. И. Черный, Р. С. Гиляревский. Информатика — новое название теории научной информации. Сб. «Научно-техн. информация», 1966, № 12, стр. 38.

мира, независимо от их национальности и отраслей науки или техники, в которых они работают. Такая замена означает не только громадное увеличение количества и разнообразия научной информации, которой обмениваются эти исследователи, но и крайнее усложнение системы передачи этой информации. Теперь для передачи научной информации нужно преодолевать не только трудности, связанные с большими расстояниями, но и границы между отраслями науки и техники, языковые барьеры и т. д.

Рис. 24. Упрощенная схема передачи научной информации в процессе развития науки

Выше уже отмечалось, что основным средством передачи научной информации во времени и пространстве был и продолжает оставаться письменный документ, имеющий самые различные формы. Такая система научной связи существует с момента возникновения науки и в целом успешно выполняет свою основную функцию. Но особенностью этой коммуникативной системы является то, что она обеспечивает доведение непосредственно до потенциального потребителя лишь небольшой доли адресованной ему научной информации. Величина этой доли особенно уменьшилась в последнее время, что объясняется несколькими причинами. Прежде всего, развитие науки и техники сопровождается быстрым ростом количества письменных документов, содержащих научную информацию по тому или иному вопросу. Например, лишь по химии и химической технологии в 1967 г. было опубликовано 310 тыс. статей, научно-технических отчетов, патентных описаний и других научных документов, причем это число ежегодно возрастает на 8,3% (период удвоения — 7,5 лет) [6]. Свыше 130 тыс. научных документов было опубликовано в 1965 г. по биологии [7]. Аналогичная картина наблюдается и в других отраслях науки и техники.

[6] Chemical literature expands. «Chemical and Engineering News», 1966, v. 44, № 23, стр. 86.

[7] P. V. Parkins. BioScience Information Service of Biological Abstracts. «Science», 1966, v. 152, № 3724, стр. 889.

Второй фактор, затрудняющий доведение письменных документов непосредственно до потенциального потребителя, — это увеличение числа языков, на которых публикуются такие документы. По данным ЮНЕСКО, в настоящее время на языках, которыми не владеет более половины ученых мира, публикуется 50% всей научно-технической литературы [8]. В одном из отчетов Национального научного фонда США указывалось, что около $^1/_3$ всей мировой научно-технической литературы сейчас издается в СССР, КНР и Японии, т. е. выходит на языках, которыми не владеет более 95% американских ученых [9]. В последнее десятилетие наблюдается также рост удельного веса научно-технической литературы, публикуемой в развивающихся странах Азии, Африки и Латинской Америки и выходящей на таких языках, как испанский, португальский, хинди, урду, арабский, суахили и т. д. Таким образом, исследователь и в дальнейшем будет все больше утрачивать возможность не только читать, но и находить в огромном потоке научно-технической литературы все документы, содержащие нужную ему научную информацию.

Наконец, возникновение у исследователя потребности в той или иной научной информации зависит от многих факторов и часто не совпадает во времени с появлением письменных документов, содержащих такую информацию. Например, русские ученые А. Г. Полотебнев и В. А. Манассеина в 1872 г. установили, что в условиях размножения зеленой плесени «не развиваются бактерии», и предложили опробовать грибок penicillium для лечения болезней. В 1929 г. это же свойство зеленой плесени было открыто А. Флемингом. Другой пример. В 1865 г. на заседании Общества естествоиспытателей в г. Брно (Чехословакия) Грегор Мендель доложил об открытии им биологического закона наследования отдельных признаков. Работа Менделя «Опыты над растительными гибридами» в том же году была опубликована в «Трудах» этого общества. Но наука того времени еще не была готова воспринять это открытие, и работа Менделя была забыта. Только в 1900 г. одновременно три исследователя — Г. де Фриз в Голландии, К. Чермак в Австрии и К. Корренс в Германии — вновь открыли этот закон.

Возникновение системы научной связи при помощи письменных документов сразу же потребовало организации промежуточных хранилищ, в которых бы накапливались и в систематизированном виде хранились письменные документы. Это давало исследователю возможность производить поиск письменных документов

---

[8] Scientific and technical translating and other aspects of the language problem. Documentation and terminology of science. Paris, Unesco, 1957, стр. 13.

[9] K. Gingold. Translations for the U. S. scientists. «Chemical and Engineering News», 1964, v. 42, N 33, стр. 88.

с интересующей его научной информацией в любое удобное для него время. Характер таких промежуточных хранилищ или библиотек изменялся в зависимости от типа накапливаемых в них документов, а также от потребностей исследователей, для обслуживания которых они создавались.

Теперь можно внести еще одно существенное уточнение в схему, которая была приведена на рис. 24. В новой схеме (рис. 25) уже имеется промежуточное хранилище, в котором исследователи отыскивают письменные документы, содержащие нужную им информацию, т. е. производят информационный поиск.

Рис. 25. Упрощенная схема циркуляции научной информации между исследователями

*Информационный поиск* — это некоторая последовательность операций, выполняемых с целью отыскания документов (статей, научно-технических отчетов, описаний к авторским свидетельствам и патентам, книг и т. д.), содержащих определенную информацию (с последующей выдачей самих документов или их копий), или с целью выдачи фактических данных, представляющих собой ответы на заданные вопросы. Информационный поиск производится при помощи *информационно-поисковых систем*. Информационно-поисковая система (ИПС) — это некоторая совокупность или комплекс связанных друг с другом отдельных частей, предназначенный для выявления в каком-либо множестве элементов информации (документов, сведений и т. д.), которые отвечают на информационный запрос, предъявленный системе. Массив элементов информации, в котором производится информационный поиск, условимся называть далее *поисковым массивом*.

ИПС подразделяются на *документальные* и *фактографические*. Документальные ИПС в ответ на вводимые в них информационные запросы выдают оригиналы, копии или адреса хранения документов, содержащих требуемую информацию. Подкласс документальных ИПС, выдающих лишь библиографические описания

искомых документов, иногда называют *библиографическими ИПС* (рис. 26).

В отличие от документальных ИПС фактографические поисковые системы предназначены для выдачи непосредственно требуемой информации (например, температуры кипения какой-либо жидкости; структурных или молекулярных формул химических соединений, обладающих некоторыми свойствами, и т. д.). Однако принципиальных различий между документальными и фактографическими ИПС нет. Если прибегнуть к сравнению,

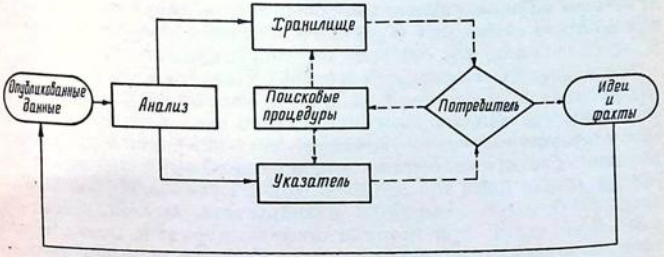



Рис. 26. Схема документальной информационно-поисковой системы

то документальные ИПС отличаются от фактографических, по-видимому, не в большей степени, чем первичные научные документы (статьи, научно-технические отчеты, монографии и т. д.), содержащие различную фактическую информацию, отличаются от справочников, в которых определенная часть этой же фактической информации представлена в систематизированном виде. Основным признаком, объединяющим документальные и фактографические ИПС в один общий класс, является то, что и документальные, и фактографические ИПС могут по запросам выдавать такую и только такую информацию, которая была ранее введена в них.

Другой, высший класс образуют *информационно-логические системы*. В отличие от ИПС информационно-логические системы должны не только выдавать по требованию ранее введенную в них информацию, но и производить, если это необходимо, логическую переработку этой информации с целью получения новой информации, которая в явном виде в них не вводилась. Информационно-логические системы могут быть практически созданы лишь после того, как будет решено множество сложнейших научных и технических задач, причем некоторые из этих задач имеют не узкоспециальный, а широкий общенаучный характер. К числу таких общенаучных задач относится, например, разработка

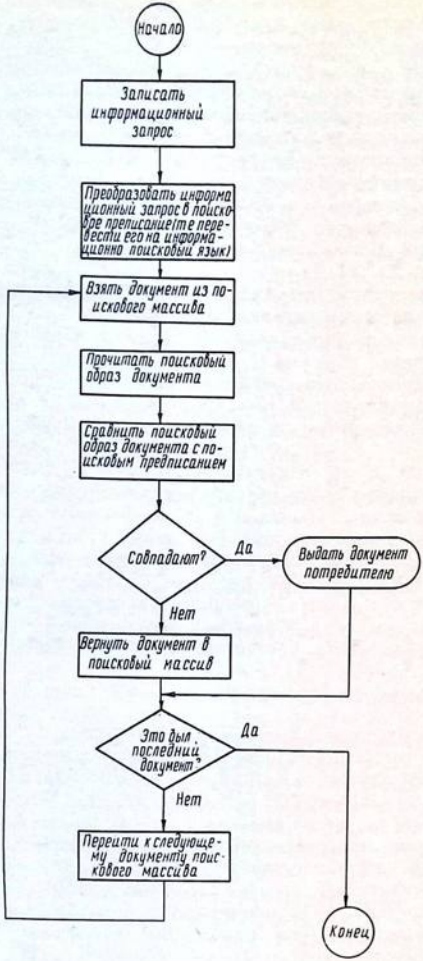



Рис. 27. Упрощенная блок-схема программы документального информационного поиска

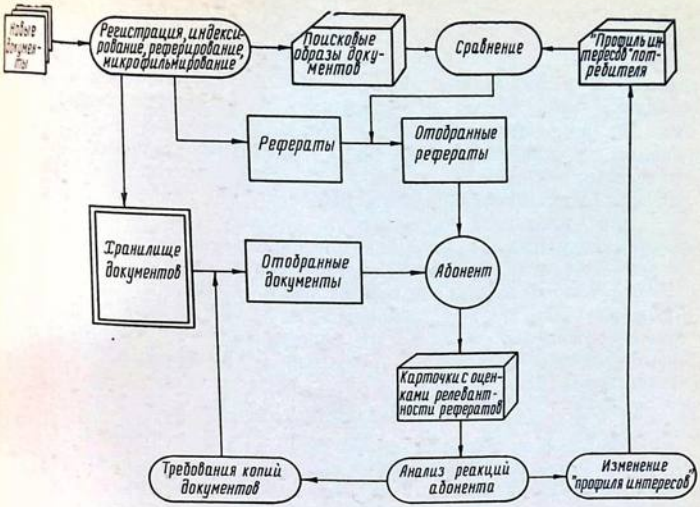



Рис. 28. Технологическая схема избирательного распределения информации

которые формулируются самими потребителями. Это особый случай информационного поиска, производимого в ограниченном массиве письменных документов. Поисковый массив образуют документы, публикуемые лишь в текущей научно-технической литературе. ИРИ в принципе мало чем отличается от ретроспективного информационного поиска. Технологическая схема ИРИ показана на рис. 28. Необходимо отметить, что системы ИРИ позволяют также находить среди обслуживаемых ими специалистов (абонентов) таких, чьи профессиональные интересы в наибольшей степени соответствуют определенной задаче. Кроме того, при их помощи можно выявлять связи новых научно-исследовательских тем с уже разрабатываемыми.

3. Поиск имен специалистов, располагающих информацией по определенному вопросу.

Для рассмотрения основных проблем документального информационного поиска уточним следующие понятия:

Под *информационно-поисковой системой* в ее абстрактном виде понимается совокупность информационно-поискового языка (с правилами перевода с естественного языка на этот язык и наоборот) и критерия смыслового соответствия между поисковыми образами документов и поисковыми предписаниями. В понятие абстрактной информационно-поисковой системы не включаются

средства ее технической реализации — каталожные карточки, перфокарты разных типов, счетно-перфорационные машины, специальные устройства и т. п. В специальной литературе информационно-поисковыми системами часто называют также специальные устройства и машины, предназначенные лишь для отыскания документов по известным адресам их хранения. В данной работе мы не относим такие устройства и машины к категории информационно-поисковых систем.

Если определить абстрактную систему как любую совокупность переменных, которые выбираются из бесконечного числа переменных, свойственных всякой реальной системе (или «машине») [11], то формально документальная ИПС (обозначим ее через $S$) может быть задана параметрами $L_r$ и $\varphi$, где $L_r$ — информационно-поисковый язык (с правилами перевода с естественного языка на этот язык и наоборот) и $\varphi$ — критерий смыслового соответствия.

$$S = f(L_r, \varphi).$$

Определим теперь эти параметры документальной ИПС. *Информационно-поисковый язык* — это определенная семантическая система, предназначенная для выражения основного смыслового содержания документов и информационных запросов с целью отыскания в массиве (собрании) таких документов, которые содержат требуемую информацию. Правила перевода с естественного языка на информационно-поисковый язык (и наоборот) обычно задаются в виде двуязычного словаря и соответствующего алгоритма. *Критерий смыслового соответствия* — это совокупность правил, по которым определяется степень смысловой близости между поисковым образом документа и поисковым предписанием. Перечисленные параметры документальной ИПС подробно рассматриваются в следующих разделах данной главы. Здесь же будут еще даны уточненные определения нескольких основных терминов, которые широко используются в последующем изложении.

*Поисковый образ документа* — это выраженное в терминах информационно-поискового языка основное смысловое содержание этого документа (а не вся информация, содержащаяся в нем!), которое поставлено в однозначное соответствие данному документу и предназначено для его отыскания в массиве других документов. Процедура, завершающаяся присвоением документу соответствующего поискового образа, называется *индексированием*. Эта процедура состоит из двух этапов. Сначала выявляется основное смысловое содержание документа (текста), т. е. на естественном языке мысленно составляется его аннотация или

[11] У. Р. Эшби. Конструкция мозга. Происхождение адаптивного поведения. М., ИЛ, 1962, стр. 40.

реферат. Затем эта аннотация или реферат по соответствующим правилам переводится на информационно-поисковый язык.

В настоящее время не известно ни одного метода, который бы позволял автоматизировать первый этап этого процесса, так как совершенно не изучен механизм его выполнения человеком. Выделение основного смыслового содержания пока может удовлетворительно решать лишь человек. Поэтому человек является непременным составным элементом любой из существующих ИПС.

*Поисковое предписание* — это выраженное в терминах информационно-поискового языка смысловое содержание информационного запроса. Проблема адекватности выражения в информационном запросе фактической информационной потребности исследователя имеет исключительно большое значение для теории информационного поиска. Поэтому эта проблема будет также рассмотрена в одном из следующих разделов данной главы. А сейчас мы переходим к обсуждению первого параметра ИПС — информационно-поискового языка.

## ИНФОРМАЦИОННО-ПОИСКОВЫЙ ЯЗЫК

*Язык* — это знаковая система любой физической природы, выполняющая познавательную и коммуникативную функцию в процессе человеческой деятельности. Язык может быть *естественным* и *искусственным*. Естественный язык — это «язык в собственном смысле, человеческий язык как орудие мысли и средство общения в отличие от его искусственных субститутов» [12]. Естественный язык функционирует не по конвенциональным правилам, которые к тому же не являются неизменными. В естественном языке нет и не может быть однозначного соответствия между словами и значениями. Как отмечал А. Тарский, «в обычной речи не существует фразы, имеющей точно определенный смысл. Едва ли можно было бы найти двух человек, которые употребляли бы каждое слово в одинаковом значении, и даже в речи одного человека значение одного и того же слова меняется в различные периоды жизни. Сверх того, значение слов повседневного языка обычно очень сложно; оно зависит не только от внешней формы слова, но также и от обстоятельств, при которых оно высказано, а иногда и от субъективно-психологических факторов» [13]. Поэтому применение естественного языка в ИПС для однозначного описания основного смыслового содержания документов и информационных запросов связано с весьма значительными трудностями.

---

[12] Философский словарь. М., Госполитиздат, 1963, стр. 535.

[13] А. Тарский. Введение в логику и методологию дедуктивных наук. М., ИЛ, 1948, стр. 59.

Вместо естественных языков в ИПС уже давно применяются специализированные искусственные языки, которые в дальнейшем мы будем называть *информационно-поисковыми языками*.

Искусственным языком называется любой вспомогательный язык, созданный людьми для каких-либо узких целей. Искусственные языки создаются и существуют лишь на базе естественных языков. Искусственный язык, который можно было бы применить в системе хранения и поиска информации, должен быть формализован. Это означает, что в нем должна быть устранена многозначность слов естественного языка и все то, что характеризует наше отношение к различным предметам, наши эмоции, волевые побуждения и т. п. (элиминация прагматической функции языка). В языке должны выражаться лишь объективные характеристики предметов и их соотношений [14]. В связи с этим необходимо отметить, что в естественном языке многозначность слов обеспечивает богатейшие возможности для вполне однозначного выражения тончайших оттенков мысли (в процессе речи эту многозначность устраняет контекст) и служит в некотором роде показателем развитости языка.

Внутреннюю структуру языка образует целостная, организованная, единая совокупность составляющих его сторон — фонетики, лексики, грамматики и словообразования, частей этих сторон и отдельных лексических единиц (фонем, морфем, слов и словосочетаний, предложений и т. п.).

*Фонетика* — это совокупность звуков и звуковых явлений языка. Искусственные языки, используемые в системах хранения и поиска информации, есть в основном языки письменные. Поэтому рассмотрение их фонетики было бы излишним. *Лексика* (или словарный состав) — это вся совокупность слов, входящих в состав какого-либо языка. Объективно существующая в языке система способов и средств построения и изменения слов и построения предложений в данном языке образует его *грамматику*, которая состоит из *морфологии* и *синтаксиса* [15]. Морфология — это совокупность объективно существующих и действующих в языке способов и средств построения и изменения слов [16]. Наименьшая (предельная, неделимая далее без потери данного качества) и регулярно воспроизводимая согласно моделям определенного языка единица системы выражения, непосредственно соотносимая с соответствующим ей элементом содержания (семемой) [17], называется *морфемой*. Морфемы делятся на *грамма-*

---

[14] Д. П. Горский. Формальная логика и язык. В кн.: «Философские вопросы современной формальной логики». М., Изд-во АН СССР, 1962, стр. 61.

[15] Б. Н. Головин. Введение в языкознание. М., «Высшая школа», 1966, стр. 23.

[16] Там же, стр. 24.

[17] О. С. Ахманова. Словарь лингвистических терминов. М., «Сов. энциклопедия», 1966, стр. 240—241.

*тические* (аффиксы) и *лексические* (лексемы). Аффиксы предназначены лишь для видоизменения значения остальной части слова (базы). Лексическая морфема — это корневая или словообразовательная морфема, выражающая основное (материальное) значение слова. Синтаксис — это совокупность объективно существующих и действующих в языке способов и средств построения предложений. Четвертую сторону языка составляет словообразование, которое определяется как совокупность способов и средств образования слов от уже существующих, а также типовых отношений, сложившихся в языке между производящими и производными словами[18].

Необходимо отметить, что приведенное выше описание структуры языка было дано в терминах традиционного языкознания. Однако в настоящее время под структурой языка часто понимают лишь совокупность фонем, морфем, конструкций и отношений, связывающих эти языковые единицы, а лексику и грамматику языка относят к другому делению языка по уровням — к уровню нормы, основными единицами которого являются слово и предложение. Определенная непоследовательность, порожденная этим отклонением, нам представляется вполне терпимой.

Перечисление типов членения языка следовало бы начать с членения на *диахронию* и *синхронию*. Диахрония — это состояние языка при переходе от одного момента к другому, т. е. динамика языка. Синхрония — состояние языка в каждый данный момент времени, т. е. статика языка.

Другое исключительно важное членение языка — *план выражения* и *план содержания*. Такое членение возникает при рассмотрении языка «с точки зрения того, как относятся его элементы к значению. План содержания, смысл, значение называется также семантикой»[19].

Следующее членение языка — *парадигматика* и *синтагматика*. Ю. С. Степанов определяет эти понятия следующим образом:

«Парадигматикой называется такой аспект рассмотрения языка, когда языковые элементы, идущие в речевой цепи по принципу «или—или», «или один—или другой», извлекаются из речевой цепи и сводятся в группы, классы. В парадигматике они расклассифицированы и сосуществуют»[20]. Синтагматика — это другой аспект языка, характеризующийся тем, что языковые элементы располагаются в речевой цепи (сообщении) только один за другим, но не одновременно; в каждый данный момент может быть высказан (использован) или один элемент,

[18] Б. Н. Головин. Введение в языкознание. М., «Высшая школа», 1966, стр. 24.
[19] Ю. С. Степанов. Основы языкознания. М., «Просвещение», 1966, стр. 6.
[20] Там же, стр. 20.

или другой, но не оба сразу [21]. Отношение парадигматики и синтагматики можно графически представить в виде двух осей, пересекающихся под прямым углом: вертикальная ось — парадигматика, а горизонтальная — синтагматика (ось времени).

Для того чтобы сообщить о чем-то, мы должны назвать интересующие нас предметы и определенным образом сочетать эти названия. Называя предмет, мы тем самым вычленяем его из множества других предметов, причем выбор названия, а следовательно в известной степени и сам характер «вычленения», зависит от тех лексико-семантических ресурсов, которыми располагает данный язык [22]. Таким образом, употребление слов в языке определяется двумя факторами: возможностью выбора слова, наиболее точно отображающего тот или иной предмет или явление объективной действительности, и возможностью сочетания этого слова с другими словами. «С одной стороны,— отмечал Ф. де Соссюр, — слова в речи, образуя цепь, вступают между собою в отношения, основанные на линейном характере языка, исключающем возможность произнесения двух элементов сразу. . . Такие сочетания, опирающиеся на протяженность, могут быть названы синтагмами. . . С другой стороны, вне процесса речи слова, имеющие между собой что-либо общее, ассоциируются в памяти так, что из них образуются группы, внутри которых обнаруживаются весьма разнообразные отношения» [23].

Слова в языке обозначают предметы и явления объективной действительности, связанные друг с другом определенными отношениями. Естественно, что такие же отношения существуют и между словами. Эти отношения обусловлены предметно-логическими, а не собственно языковыми (семантическими) факторами, т. е. относятся к категории внеязыковых связей. Благодаря таким предметно-логическим связям слова на основе того или иного семантического признака объединяются в лексико-семантические группы, которые называются *парадигмами*.

В результате объединения и противопоставления слов по различным признакам образуются ряды связанных между собой лексико-семантических парадигм, которые можно назвать *тематическими группами*. Например, слова *береза*, *гриб*, *ель*, *кедр*, *куст*, *лавр*, *самшит*, *туя*, *цветок* и т. д. образуют тематическую группу «растения». Но внутри этой группы перечисленные слова объединяются в лексико-семантические парадигмы «деревья»: «хвойные деревья», «вечнозеленые деревья» и т. д., причем парадигме «деревья» подчинены пересекающиеся парадигмы «хвой-

[21] Ю. С. Степанов. Основы языкознания. М., «Просвещение», 1966, стр. 17.

[22] См.: Д. Н. Шмелев. Очерки по семасиологии русского языка. М., «Просвещение», 1964, стр. 128.

[23] Ф. де Соссюр. Курс общей лингвистики. М., Соцэкгиз, 1933, стр. 121.

ные деревья» (*ель+кедр+туя*) и «вечнозеленые деревья» (*ель+кедр+лавр+самшит+туя*). Из приведенного примера следует, что парадигматические отношения в лексике не только многоступенчаты, но и многомерны (неоднолинейны). Одно и то же слово может быть одновременно членом нескольких лексико-семантических парадигм, т. е. входить в различные ряды, в которых слова противопоставлены друг другу по какому-то определенному семантическому признаку.

Итак, объединение слов в лексико-семантические парадигмы происходит на основании предметно-логических (родо-видовых и других ассоциативных) связей, существующих между предметами и явлениями объективной действительности, которые обозначены этими словами. Парадигматические отношения предопределяют отбор слов для какого-либо сообщения, но сами остаются за его пределами.

Другой тип отношений между словами — *синтагматические отношения*, в которые слова вступают в пределах конкретного сообщения, фразы. «Объединение слов в тексте, — пишет Д. Н. Шмелев, — подчинено существующим в языке синтаксическим правилам. В то же время сочетаемость каждого отдельного слова в пределах допускаемых последними моделей обусловлена его собственной индивидуальной семантикой»[24].

Из всего сказанного выше следует, что парадигматика опирается на определенную совокупность или инвентарь связанных между собой языковых единиц, а синтагматика требует определенных правил соединения этих единиц в речевой цепи (сообщении). Поэтому парадигматику любого языка можно с определенными оговорками и ограничениями отождествить с лексикой этого языка, а его синтагматику — с грамматикой[25].

Теперь можно попытаться дать определение информационно-поискового языка. *Информационно-поисковый язык* — это специализированный искусственный язык, предназначенный для выражения основного смыслового содержания документов и/или только информационных запросов с целью отыскания в некотором множестве документов таких, которые отвечают на поставленный информационный запрос[26]. Для того чтобы правильно сформу-

[24] Д. Н. Шмелев. Очерки по семасиологии русского языка. М., «Просвещение», 1964, стр. 188.

[25] См.: Ю. С. Степанов. Основы языкознания. М., «Просвещение», 1966, стр. 6; Р. А. Будагов. Введение в науку о языке. М., «Просвещение», 1965, стр. 228.

[26] Из приведенного выше определения следует, что информационно-поисковый язык может использоваться для выражения основного смыслового содержания не документов, а только информационных запросов. Такой подход применен, например, в документальной ИПС, которая создана в Научно-исследовательском центре им. Томаса Уотсона фирмы «Интернэшнал бизнесс машинз, корп.» (США). В качестве поисковых образов документов в этой ИПС используются их аннотации и рефераты на естественном языке, а поис-

лировать основные требования, которые предъявляются к информационно-поисковому языку (ИПЯ), на наш взгляд, целесообразно привести несколько определений ИПЯ в формулировке разных специалистов.

1. «Информационный язык является средством для компактной, строго однозначной записи содержания документа в форме, приемлемой для машинного поиска информации. . . Необходимо подчеркнуть разницу между понятиями «информационного языка» и «формализованного языка» (к последнему близко по содержанию понятие формальной системы). Хотя всякий информационный язык включает в себя элементы формализации, он не строится вполне формально, т. е. в нем нет аксиоматики и строго фиксированных правил формально-логического вывода. . . Основное требование, предъявляемое к информационному языку, сводится к тому, чтобы он давал возможность выражать существенные стороны содержания индексируемого и кодируемого документа в форме, пригодной для поиска при помощи машины. При кодировании содержания документа важно провести эту операцию так, чтобы при поисках информации наиболее существенные сведения, содержащиеся в документе, находились легче, чем второстепенные» [27].

2. Информационный язык (ИЯ) — это искусственный язык, который «должен быть лишен семантической неоднозначности реального языка. Каждому понятию ИЯ, в частности объектам или виду отношений между ними, должна соответствовать только одна строго определенная последовательность символов, и, наоборот, каждая последовательность символов должна обозначать только одно понятие. Грамматика ИЯ должна быть построена строго формально. Из элементарных символов исходных понятий составляются слова, из слов — фразы ИЯ. Выражения, переведенные на ИЯ, должны сохранять смысл исходных» [28].

3. «Информационный язык — искусственный язык, создаваемый для записи научно-технических сведений, помещаемых в ЗУ информационной машины или системы. Ни один из естественных языков, по-видимому, не может быть применен в качестве информационного. Этому препятствует их недостаточно

---

ковые предписания формулируются в виде перечней соответствующих слов и фраз естественного языка, логических связок, а также количественных показателей, устанавливающих степень необходимого и достаточного совпадения поисковых образов документов с поисковым предписанием. (См.: J. J. Magnino. Jr. IBM Technical information retrieval center — normal text techniques. ITIRC-002, Yorktown Heights, N. Y., Thomas J. Watson Research Center. 1965).

[27] Г. Э. Влэдуц, В. В. Налимов, Н. И. Стяжкин. Научная и техническая информация как одна из задач кибернетики. «Успехи физ. наук», 1959, т. 59, вып. 1, стр. 41—42.

[28] Г. Влэдуц, А. Л. Сейфер. Машинные информационно-логические методы в химии. «Заводская лаборатория», 1962, т. 28, № 10, стр. 1225.

четкая структура (большое число грамматических исключений, неясных случаев и т. п.) и их неоднозначность (например, явление омонимии и синонимии, многозначности слов). Логические связи во фразах реальных языков также выражены недостаточно отчетливо. . . Информационный язык должен быть приспособлен для использования его в машине. Производя различные действия над текстами, помещенными в ЗУ, машина не может обращаться к их интуитивно известному смыслу, но должна ограничиваться лишь анализом их формальной структуры. Поэтому информационный язык должен быть однозначным (каждое его выражение должно допускать лишь одно истолкование), удобным для отождествления различным образом записанных фактов и для производства логических выводов» [29].

4. «Язык индексирования (index language) мы определяем как совокупность используемых в информационно-поисковой системе индексных терминов (indexing terms), правил для их использования и любых вспомогательных средств, например перекрестных ссылок, классификационных таблиц или входного (lead-in) словаря» [30].

5. «Язык индексирования (index language) — это система символов, предназначенная для выражения предметного содержания документов и являющаяся отдельной составной частью и переменной в информационно-поисковой системе» [31].

6. «Информационно-поисковый язык (retrieval language) — это искусственный язык, словами которого являются номера (numbers) систематизированных понятий, а синтаксис отражает семантические отношения между этими понятиями» [32].

7. «Документальный язык (langage documentaire) — это система символов, используемых при индексировании документов. Такая система не является самостоятельным языком в строгом значении этого термина; выражение «документальный язык» основано лишь на противопоставлении фактов, на сравнении обычно сжатого словарного состава и/или грамматики, употребляемых для построения поисковых образов (indexées), с богатством естественного языка, который служит основой для создания документального языка. В этом смысле документальный язык

---

[29] В. А. Козмидиади. Информационный язык. «Автоматизация производства и промышленная электроника», т. 1. М., «Сов. энциклопедия», 1962, стр. 491.

[30] C. W. Cleverdon. The testing and evaluation of the operating efficiency of the intellectual stages of information retrieval systems. В кн.: «P. Atherton (ed.) Classification research. Proceedings of the 2 nd international study conference (Elsinore, Denmark, 14—18th Sept. 1964)». Copenhagen, Munksgaard, 1965, стр. 445.

[31] J. Melton. A use for the techniques of structural linguistics in documentation research. Там же, стр. 466.

[32] Current research and development in scientific documentation, N 11 (NSF-63-5). Washington, U. S. Government printing office, 1963, стр. 377.

иногда называют искусственным языком и рассматривают в двух традиционных аспектах: семантическом и грамматическом. В широком смысле правила и символы, применяемые для работы с выражениями на документальном языке, считаются принадлежащими к этому же языку» [33].

8. «Документальный язык (langage documentaire) — это средство связи, служащее для общения между одними людьми и другими, являющимися авторами документов. Иначе говоря, документальным языком называется такой язык, которым пользуется человек, когда он хочет найти определенные документы, хранящиеся среди других документов в определяемом разными способами месте. Документальный язык обычно бывает письменным (*écrit*) и может выражаться в виде диаграммы. Это не обязательно язык (*langue*) [в собственном смысле этого слова]» [34].

В некоторых определениях (1, 3) указано, что ИПЯ должен быть пригодным для механизации информационного поиска. С этим требованием нельзя согласиться, так как имеются ИПЯ (например, библиотечно-библиографические и алфавитно-предметные классификации), которые предназначены для использования в немеханизированных документальных ИПС. Кроме того, из определений (1, 3) не ясен смысл выражений «поиск при помощи машины» и «использование в машине». Если в данном случае имеется в виду лишь механизация процедуры формального сравнения поисковых образов документов с поисковым предписанием, то это в принципе возможно при использовании любого ИПЯ. Если же речь идет о выделении из неадаптированных полных текстов документов, написанных на естественном языке, их основного смыслового содержания и выражения этого содержания в терминах ИПЯ, то такая процедура при существующем уровне наших знаний о внутренних закономерностях человеческого мышления не может быть механизирована. Неясно также, в каком смысле здесь используется термин «машина». Ведь машиной может быть названа любая целесообразная система, в том числе и ящик с каталожными карточками.

В связи с этим необходимо отметить, что до настоящего времени все еще наблюдается определенная фетишизация применения в области документального информационного поиска специальных устройств, СПМ и ЭЦВМ. Это явление — своеобразная «детская болезнь», свойственная начальному периоду исследований в области механизации информационного поиска. Оно объясняется недостаточным пониманием действительных проблем информационного поиска, решение которых невозможно без глу-

[33] R. C. Cros, J.-C. Gardin, F. Lévy. L'Automatisation des recherches documentaires. Un modèle général le Syntol. Paris, Gauthier-Villars, 1964, стр. 243.

[34] M. Coyaud. Introduction a l'étude des langages documentaires. Paris, Klincksiek, 1966, стр. 16.

бокого изучения механизмов человеческого мышления. Нечто подобное наблюдалось ранее в области машинного перевода.

Механизация ради механизации никому не нужна. Эта истина тривиальна, но ее часто игнорируют. Механизация и автоматизация документального информационного поиска целесообразна лишь тогда, когда она позволяет удовлетворить хотя бы одно из следующих требований: а) получить реальный экономический эффект (это означает, что стоимость машинного поиска документов — с учетом всех затрат человеческого труда и стоимости используемого оборудования — должна быть ниже, чем при использовании немеханизированных ИПС); б) сократить до минимума время информационного поиска, когда решающую роль играет фактор времени, а не стоимость; в) решать такие информационно-поисковые задачи, которые в принципе не могут быть решены с помощью немеханизированных ИПС.

Некоторые авторы [35], вместо терминов «информационный язык», «информационно-поисковый язык», «документальный язык» и «язык индексирования» (или наряду с этими терминами), используют также термин «машинный язык». Термин «машинный язык» нельзя признать удачным, так как в научно-технической литературе он употребляется по крайней мере в трех значениях:

—язык для поиска информации;

—язык, который используется для записи информации в форме, воспринимаемой той или иной машиной (обычно в качестве такого языка используется двоичная система счисления);

—«формальный язык для описания алгоритмов решения задач, учитывающий, что эти алгоритмы должны быть реализованы на вычислительной машине» [36].

Применение одного и того же термина для выражения совершенно разных понятий может создать ложное впечатление о тождественности этих понятий и стать источником серьезных недоразумений. Например, в «Словаре лингвистических терминов» О. С. Ахмановой (М., «Сов. энциклопедия», 1966, стр. 534) дано следующее определение информационного языка: «Я з ы к и н ф о р м а ц и о н н ы й *англ.* machine language. Средство, обеспечивающее компактную и строго однозначную запись со-

---

[35] J. W. P e r r y, A. K e n t, M. B e r r y. Machine literature searching. New York, Interscience Publishers, 1956, XI, 166 стр.; J. P e r r y, A. K e n t. Documentation and information retrieval. An introduction to basic principles and cost analysis. Cleveland, Ohio, Western Reserve University Press, 1957, XII, 156 стр.; Л. И. Г у т е н м а х е р. Электронные информационно-логические машины. Изд. 2. М., Изд. АН СССР, 1962; В. В. И в а н о в. Лингвистические вопросы создания машинного языка для информационной машины. В сб: «Материалы по машинному переводу». Л., Изд-во ЛГУ, 1958, стр. 10—39 и др.

[36] Э. З. Л ю б и м с к и й. Машинный язык. «Автоматизация производства и промышленная электроника», т. 2. М., «Сов. энциклопедия», 1963, стр. 260.

держания в форме, приемлемой для непосредственной подачи на механизмы (электронные, счетно-аналитические и т. п.)». Фактически здесь дано определение не информационного, а машинного языка в его втором значении. Если исходить из этого определения, то тогда одно из основных требований к ИПЯ состояло бы в том, чтобы записи на таком языке могли восприниматься машиной. При таком подходе ИПЯ начинает рассматриваться в основном как система не смыслового, а физического кодирования информации с целью ее обработки в машине. Однако кодирование сообщений на ИПЯ, т. е. соответствующее изменение их физической формы, не имеет прямого отношения к самому ИПЯ.

Кроме того, в термине «машинный язык», употребляемом в значении «информационно-поисковый язык», на передний план выдвигается требование машинной реализации информационного поиска. По-видимому, на этом основании О. С. Ахманова в своем предисловии к книге П. Н. Денисова «Принципы моделирования языка» отграничивает языки для поиска информации от одного из основных типов таких языков — библиотечно-библиографических классификаций [37]. О неправомерности требования, чтобы информационный поиск производился обязательно при помощи какой-либо машины, а также о неопределенности самого термина «машина» уже говорилось выше.

Возражение вызывает содержащееся в определении (3) требование, чтобы ИПЯ был пригоден для производства логического вывода. Это требование объясняется расширительным толкованием термина «информационный язык», который в данном случае понимается как язык, используемый в информационно-логических системах. Логический вывод, как известно, возможен лишь в рамках формальных систем. Между тем в определении (1) было справедливо отмечено, что ИПЯ не является формальной системой в строгом понимании этого термина, хотя он может и должен содержать в себе определенные элементы формализации. Ниже приводится определение формальной системы, называемой также абстрактным языком, семантической системой, логическим исчислением и т. д. «Абстрактный язык, или формальная семантическая система, состоит из списка элементарных символов (знаков), правил образования (устанавливающих, какие комбинации знаков допускаются), правил преобразования (устанавливающих, какие допускаются преобразования выражений с целью получения логического вывода) и правил интерпретации (устанавливающих, какой смысл надлежит приписывать выражениям, составленным по правилам образования)» [38].

[37] П. Н. Денисов. Принципы моделирования языка. М., Изд-во МГУ, 1965, стр. 4.

[38] В. А. Успенский. К проблеме построения машинного языка для информационной машины. «Проблемы кибернетики», 1959, вып. 2, стр. 45.

Некоторые ученые считают, что язык как система содержательных символов вообще не может быть сведен к формальной системе без разрушения его истинной природы. По мнению М. Таубе, именно невозможностью полной формализации языка объясняется то, что исторически грамматика явилась нормативом, а не описательной наукой [39].

Если из числа составных частей абстрактного языка исключить правила, которые устанавливают допустимые в нем преобразования с целью получения логического вывода, то ИПЯ можно определить как специализированную семантическую систему, состоящую из списка используемых в ней элементарных символов (алфавита), правил образования и правил интерпретации. В качестве алфавита, понимаемого в данном случае как любая конечная совокупность знаков (букв), в разных ИПЯ используются различные наборы элементарных символов. Например, в Универсальной десятичной классификации (УДК) для этой цели используются следующие символы: 0; 1; 2; 3; 4; 5; 6; 7; 8; 9; -[дефис]; . [точка]; + ; : ; / ; =; ' [апостроф]; « » [кавычки]; ( ) [круглые скобки], а также все строчные и прописные буквы соответствующего естественного языка, специальные знаки и символы и т. д. (применяются для словесного выражения дополняющих индексы УДК имен собственных, предметных заголовков, химических формул и т. п.) [40]. В языке системы «Пусто-Непусто-4» применяются два алфавита: все арабские цифры и все буквы русского алфавита [41]. Алфавит Семантического кода (США) состоит из следующих символов: всех прописных букв латинского алфавита, всех арабских цифр, а также знаков *; , [запятая]; - [дефис]; &; ¶[; . [точка]; /; □ и $ [42].

При построении ИПЯ выбор алфавита определяется в основном лишь соображениями практического удобства. Однако это не означает, что другие стороны ИПЯ вообще никак не связаны с его алфавитом и не устанавливают никаких ограничений на его выбор. Ясно, что выбор алфавита для ИПЯ в какой-то мере определяется внутренней структурой самого ИПЯ, предполагаемыми схемами и средствами технической реализации ИПС и другими факторами.

---

[39] М. Таубе. Вычислительные машины и здравый смысл. Миф о думающих машинах. М., «Прогресс», 1964, стр. 18.

[40] П. Герман. Практическое применение Универсальной десятичной классификации. Техника классифицирования. М., ВИНИТИ, 1964, стр. 13. Квадратные скобки, в которых здесь и далее дается расшифровка символов, а также точка с запятой, которыми они разделяются, не входят в алфавит соответствующего ИПЯ.

[41] Э. Бернштейн, Д. Лахути, В. Чернявский. Вопросы теории поисковых систем. М., ВНИИЭМ, 1966, стр. 64.

[42] J. W. Perry, A. Kent. Tools for machine literature searching Semantic code dictionary. Equipment. Procedures. New York, Interscience Publishers, Inc., 1958.

Правила образования (построения) делятся на *морфологические* и *синтаксические*. Морфологические правила определяют процедуру построения слов ИПЯ из его морфем, а синтаксические — процедуру построения в данном ИПЯ предложений (фраз) из этих слов. При этом необходимо подчеркнуть, что слова в ИПЯ строятся не из символов алфавита, а из морфем. В структуру ИПЯ непосредственно не входят фонемы — кратчайшие единицы, различающие слова по их звучанию. Поэтому символы алфавита в ИПЯ не могут обозначать никаких меньших структурных элементов языка, чем морфемы [43]. Имеющие смысл отдельные алфавитные символы ИПЯ и их наименьшие сочетания можно рассматривать как своеобразные *лограммы* — знаки, обозначающие содержание, значение слов. Под словом в данном случае понимается н а з в а н и е, и м я вещи или явления, тогда как морфема может быть лишь сегментом такого имени. Поэтому слово может состоять как из одной, так и из нескольких морфем. К категории слов ИПЯ относятся также лексикализированные словосочетания, т. е. такие, которые функционируют в качестве эквивалентов отдельных слов.

В Семантическом коде морфемами являются семантические факторы, инфиксы и числовые суффиксы. В УДК морфемами могут быть только целые индексы из основной и вспомогательных таблиц. В индексе УДК самостоятельное значение имеет лишь первый символ, а значение каждого последующего символа в этом индексе определяется всеми предшествующими символами. Например, в индексе 025.13 (Репродуцирование и поиск документов) только первая цифра 0 имеет самостоятельное значение (0-Общий отдел). Цифра 2 в этом индексе имеет значение лишь в сочетании с цифрой 0 (02-Библиотечное дело. Библиотековедение); цифра 5 — лишь в сочетании с цифрами 02 (025-Работа библиотеки: отделы. Службы и различные виды работы) и т. д.

В некоторых ИПЯ словарный состав представляет собой фрагмент лексики естественного языка, на базе которого строится этот ИПЯ. Очевидно, что такие ИПЯ не имеют морфологии.

Синтаксические правила — обязательный элемент любого ИПЯ. В некоторых ИПЯ для соединения слов в предложения (фразы) применяются специальные лексические средства. Например, в УДК роль таких средств выполняют символы +; /; ' [апостроф]; :; =; ( ); « »; .00 [точка нуль нуль+цифра]. В Семантическом коде для этой цели используются символы *; , [запятая]; —; & и ⅂⌊.

Но существуют и такие ИПЯ, в которых для соединения слов в предложения применяется лишь их координация (соподчинение) и/или подчинение.

[43] Сказанное выше не относится к ИПЯ, в которых основной словарный состав представляет собой фрагмент лексики соответствующего естественного языка.

Последний элемент ИПЯ, если его рассматривать как специализированный абстрактный язык, — это правила интерпретации, т. е. правила перевода слов и предложений (фраз) ИПЯ на соответствующий естественный язык. Эти правила задаются в виде двуязычных словарей, в которых каждой лексической единице ИПЯ ставится в соответствие определенное слово или выражение естественного языка, и наоборот. В такой словарь включаются также все символы, применяемые в данном ИПЯ для соединения слов в предложения (фразы). Кроме того, правила интерпретации для ИПЯ, как и правила построения, формулируются на естественном языке в специальных инструкциях, методиках и т. д.

Сформулируем теперь основные требования, предъявляемые к любому ИПЯ. Эти требования состоят в следующем.

1. В ИПЯ должна быть полностью устранена семантическая неоднозначность словарного состава (лексики). Следовательно, в ИПЯ каждое понятие должно выражаться одним и только одним словом, т. е. одной строго определенной последовательностью символов, и наоборот, каждое слово должно выражать одно и только одно понятие.

2. Грамматика ИПЯ должна быть построена строго формально. Каждое выражение на ИПЯ должно допускать лишь одно истолкование.

3. Из лексики и грамматики ИПЯ должны быть устранены все элементы, как-то связывающие сообщение с автором или адресатом этого сообщения (элиминация прагматического плана языка).

Лексические элементы в ИПЯ выступают как единицы, в значительной мере лишенные лексических значений, т. е. таких качеств, как эмоциональность и экспрессивность, полисемия, связь со значением и т. д. Эти качества (характеристики) оказываются «избыточными» для лексических элементов ИПЯ, где за ними сохраняется только чисто логическое или предметно-понятийное содержание [44].

Как любой другой искусственный язык, ИПЯ строится на базе соответствующего естественного языка, из которого приходится многое исключить. По отношению к естественному языку ИПЯ можно рассматривать как определенный *метаязык*, так как объектом ИПЯ являются содержание и выражения естественного языка. ИПЯ обладает соответствующей грамматикой, в которой может отсутствовать морфология, но обязательно должен содержаться синтаксис, т. е. правила образования (построения) «осмысленных» выражений из лексических единиц ИПЯ. Сами же лексические единицы могут быть в готовом виде отобраны для

[44] В. А. Звегинцев. Очерки по общему языкознанию. М., Изд-во МГУ, 1962, стр. 40—53.

ИПЯ из словарного состава соответствующего естественного языка [45].

Весьма желательно также обозначение в лексике ИПЯ парадигматических связей между словами. Авторы статей и книг по информационному поиску часто называют такие связи *аналитическими* и реже — *базисными*. «При разработке машинного языка, — писали Дж. Перри, А. Кент и М. Берри в своей книге «Поиск литературы с помощью машин», — должна иметься в виду не только необходимость записи (в ясной форме) аспектов предметного содержания, позволяющей путем автоматических операций идентифицировать нужные (pertinent) документы и графические произведения (graphic records). Нужно также помнить о необходимости отсеивания документов и графических произведений, которые не представляют интереса с точки зрения данной проблемы или ситуации, сложившейся перед потребителем (inquirer). Для удовлетворения этого последнего требования особенно выгодным может оказаться обозначение синтетических отношений. С другой стороны, при построении машинного языка весьма желательно ясное указание аналитических связей, т. к. это может дать опорные точки (reference points) для позитивной идентификации документов и графических произведений» [46]. *Синтетическими* в данном случае названы синтагматические отношения, иногда именуемые также *текстуальными*.

Еще одно весьма желательное свойство ИПЯ — это возможность формализации процедуры перевода с естественного языка на ИПЯ (и наоборот), а также процедуры установления смысловой эквивалентности (частичной или полной) любых двух выражений на ИПЯ. Под формализацией в данном случае понимается разложение указанных процедур на их составные элементы и последовательное описание этих элементов.

Для сравнения эффективности различных типов ИПЯ введем понятие *семантической силы* искусственного языка, под которым будем понимать способность искусственного языка точно и полно выражать (передавать) смысл любого сообщения. Очевидно, что семантическая сила ИПЯ будет тем больше, чем выше степень развития его парадигматики и синтагматики.

Необходимо подчеркнуть, что семантическая сила — понятие сугубо относительное и применимо только к искусственному языку. Оно показывает, насколько данный искусственный язык в своей коммуникативной функции уступает естественному языку, на базе которого он построен. По мере развития естественного языка соответственно развиваются — хотя и с определенным

[45] В. В. Иванов. Лингвистические вопросы создания машинного языка для информационной машины. В кн.: «Материалы по машинному переводу». Сб. 1. Л., Изд-во ЛГУ, 1958, стр. 13, 24.

[46] J. W. Perry, A. Kent, M. Berry. Machine literature searching. New York, Interscience Publishers, 1956, стр. 69.

отставанием — все ИПЯ, построенные на его базе. Таким образом, при оценке семантической силы любого ИПЯ имеет смысл рассматривать этот ИПЯ не только в его синхронии, но и в диахронии.

В настоящее время для целей информационного поиска документов используются три основных типа ИПЯ:

—библиотечно-библиографические классификации;

—алфавитно-предметные классификации [47];

—языки дескрипторного типа.

Для сравнительной оценки семантической силы разных ИПЯ полезно рассмотреть особенности словарного состава (лексики) таких языков.

Существуют такие ИПЯ, в которых словосочетания и фразы, выражающие сложные понятия (т. е. состоящие из двух и более простых понятий), заданы в словаре, наряду со словами, выражающими простые понятия. Это означает, что в таких лексикализированных словосочетаниях или фразах образующие их слова связаны координатной (соподчинительной) или какой-либо иной связью до индексирования. ИПЯ такого типа называются *предкоординированными* (pre-coordinate). Условимся в дальнейшем обозначать такие языки символом $L_1$. Словарный состав языков $L_1$ в какой-то мере напоминает двуязычные разговорники, в которых заранее составлены наиболее употребительные, с точки зрения составителей, фразы. Когда применяется язык $L_1$, то для обозначения основного смыслового содержания документов можно использовать только отдельные, никак не связываемые между собой слова, словосочетания и фразы, взятые из словарного состава данного языка. Таким образом, при помощи языка $L_1$ производится классификация документов, т. е. их отнесение к классам, обозначенным лексическими единицами этого языка. Семантическая сила языка $L_1$ определяется главным образом тем, насколько тщательно и исчерпывающе разработана его лексика. При составлении словаря для языка $L_1$, которое производится до его использования, недостаточно учесть лишь специфику информационных запросов прошлого и настоящего. Необходимо также предвидеть изменения информационных потребностей ученых и инженеров в будущем. Это очень трудная задача, удовлетворительное решение которой едва ли возможно. Из сказанного следует, что языкам $L_1$ органически присуща недостаточность словарного состава, а следовательно, и небольшая семантическая сила. К языкам $L_1$ относятся иерархические и алфавитно-предметные классификации.

---

[47] Строго говоря, алфавитно-предметные классификации являются разновидностью библиотечно-библиографических классификаций. Однако в библиографии и библиотековедении под библиотечно-библиографическими классификациями понимаются лишь иерархические и, по-видимому, фасетные классификации. Авторы данной работы не видят большого смысла в изменении этой сложившейся терминологии.

Другой тип ИПЯ — это такие, в которых лексические единицы (слова) объединяются в «предложения» (поисковые образы) лишь во время индексирования документов или даже в процессе их поиска. Такие ИПЯ называются *посткоординируемыми* (post-coordinate). Будем обозначать их символом $L_2$. До индексирования лексические единицы языков $L_2$ не связаны никакими синтагматическими отношениями. Можно провести определенную аналогию между языками $L_2$ и алфавитом любого естественного языка. В естественном языке соединение букв в слова производится лишь в процессе письма, а до этого буквы никак не связаны одна с другой. Из 33 букв русского алфавита можно составить практически бесконечное число разных слов. Точно так же из сравнительно небольшого числа лексических единиц языка $L_2$ можно построить «предложение» (поисковый образ), выражающее практически любой смысл. Очевидно, что языки $L_2$ будут семантически более сильными, чем языки $L_1$. Ниже приводится классификация ИПЯ с точки зрения их принадлежности к языкам $L_1$ и $L_2$ [48].

1. Языки $L_1$
1. 1. Иерархические библиотечно-библиографические классификации.
1. 2. Фасетные классификации (с фиксированной последовательностью фасетов в фасетной формуле).
1. 3. Фасетные классификации (с возможностью перестановок в фасетной формуле).
1. 4. Алфавитно-предметные классификации.
2. Языки $L_2$
2. 1. Посткоординируемые фасетные классификации.
2. 2. Языки дескрипторного типа.

Отсюда видно, что библиотечно-библиографические и алфавитно-предметные классификации принадлежат к языкам $L_1$ (в литературе была описана лишь одна посткоординируемая фасетная классификация (2.1), относящаяся к языкам $L_2$) [49]. Таким образом, главная особенность библиотечно-библиографических и алфавитно-предметных классификаций состоит в том, что их словарный состав задается в виде более или менее фиксированного списка слов, словосочетаний и фраз. При переводе текста, выражающего основное смысловое содержание документа (такой текст может лишь подразумеваться) или информационного запроса с естественного языка на язык $L_1$, можно пользоваться только словами, словосочетаниями и фразами, содержащимися в фиксированном списке. Введение в языки $L_1$ новых лексических единиц строго ограничено; оно возможно лишь до индексирования документов, т. е. при создании языка. Весьма ограни-

[48] D. J. Campbell. Making your own indexing system in science and technology (clasification and keyword systems). «Aslib Proceedings», 1963, v. 15, № 10, стр. 286.

[49] См.: J. S. Rosser. Co-ordinate indexing in an aeronautical library. «Library Association Record», 1958, v. 60, № 4, стр. 117—119.

чены также возможности построения из лексических единиц таких языков правильных (осмысленных) выражений. Поэтому библиотечно-библиографические и алфавитно-предметные классификации не обладают такой «семантической силой», которая необходима для эффективного информационного поиска по запросам любого типа.

В связи с этим необходимо подчеркнуть, что основные трудности информационного поиска возникают вследствие того, что процессы индексирования документов и их поиска разделены во времени, причем иногда весьма значительно. Если бы можно было заранее предвидеть все варианты информационных запросов, которые могут возникнуть у потребителей информации, то документы можно было бы адекватно заиндексировать, по-видимому, средствами любого ИПЯ. Возможно, что для этого пришлось бы в необходимой степени расширить основной словарный состав библиотечно-библиографической или алфавитно-предметной классификации. Но такое условие не выполнимо в принципе: специфика информационных запросов зависит от развития науки и техники, от изменяющихся научных интересов потребителей информации, их индивидуальных особенностей и т. д., т. е. определяется факторами, которые фактически не поддаются учету и предвидению.

Возможна также следующая классификация ИПЯ по типам организации их словарного состава, который в данном случае определяется как упорядоченное или неупорядоченное множество знаков (слов естественного языка, цифровых и буквенно-цифровых индексов и т. д.), предназначенных для обозначения (выражения) основного смыслового содержания документов (см. схему, приведенную на рис. 29) [50]. В этой схеме естественные языки и ИПЯ обозначены соответственно символами $L_n$ и $L_r$. Эксплицитными словарями типа $L_n \rightarrow L_r$ названы такие, в которых выражены парадигматические связи между лексическими единицами $L_r$. В имплицитных словарях такие связи не обозначены. Словари типа $L_n \rightarrow L_r$ предназначены для перевода с $L_n$ на $L_r$. «Положительными» словарями названы собственно словари лексических единиц $L_r$, а «отрицательными» словарями — подмножества слов естественного языка, не входящих в словарный состав $L_r$ (т. е. дополняющих его до словарного состава $L_n$, на базе которого построен данный $L_n$). Примером «отрицательных» словарей могут служить списки неключевых слов, используемые при составлении на ЭЦВМ пермутационных указателей.

[50] J.-C. Gardin. Free classifications and facet classification. Their exploitation with computers. В кн.: P. Atherton (ed.). Classification research. Copenhagen, 1965, стр. 161—168; J.-C. Gardin. Éléments d'un modèle pour la description des lexiques documentaires. «Bull. des bibliothèques de France», 1966, v. 11, № 5, стр. 171—182.

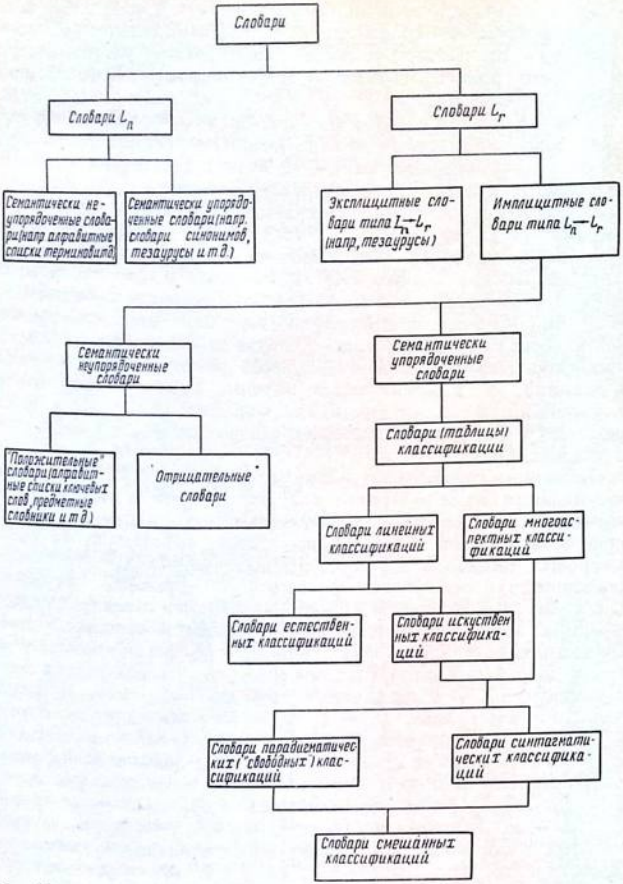



Рис. 29. Классификация информационно-поисковых языков по типам организации их словарного состава

В приведенной схеме применено еще два термина, нуждающихся в пояснениях. Под парадигматическими классификациями (Ж.-К. Гардэн называет их также «свободными») понимаются такие, в которых положение термина в таблице (схеме) не предопределяет его логической функции в «предложении» (поисковом образе или предписании). Для обозначения логической функции данного термина в каком-либо «предложении» нужно снабдить этот термин соответствующим синтаксическим определителем. В синтагматических классификациях основным критерием синтаксической организации терминов является их принадлежность к разным семантическим категориям. Единственный способ выражения логических связей между такими терминами в «предложении» — это порядок их следования в этом «предложении». В качестве примера синтагматической классификации Ж.-К. Гардэн называет фасетные классификации с фиксированным порядком фасетов в формуле.

Для единообразного анализа и сравнения структуры различных ИПЯ Дж. Мелтон [51] предложила использовать определительную модель языка, построенную П. Гарвиным (США) [52]. Эта модель, как и все классические формы определений, строится таким образом, чтобы содержание определяемого понятия «язык» раскрывалось через ближайшее родовое понятие (genus proximum) и видовое отличие (differentia spezifica).

В модели П. Гарвина ближайшее родовое понятие по отношению к видовому понятию «язык» есть понятие «знаковая система». Знаки в системе определенным образом связаны друг с другом. Поэтому для описания какого-либо языка недостаточно просто перечислить все его элементы, а необходимо еще охарактеризовать связи между ними.

Вторая часть определения языка гласит, что язык — это знаковая система, структура которой имеет три группы уровней (levels). П. Гарвин определяет уровень как множество структурных связей, относящихся к какому-либо количественному аспекту или свойству языка, и выделяет в языке три следующих «измерения»:

1) типы языковых единиц;
2) степень сложности этих единиц;
3) характер отношений между этими единицами.

Два низших уровня, которые П. Гарвин называет *уровнями построения* (lavels of structuring), оперируют с единицами языка. Языковые единицы могут быть двух видов: а) единицы, имеющие

[51] J. Melton. A use for the techniques of structural linguistics in documentation research. В кн.: P. Atherton (ed.). Classification research. Copenhagen, 1965, стр. 466—480.

[52] P. L. Garvin. The definitional model of language. В кн.: P. L. Garvin (ed.). Natural language and computer. New York, McGraw-Hill, 1963, стр. 3—22.

форму и значение (морфемы), и б) единицы, не имеющие самостоятельного значения и используемые лишь для построения или различения морфем (фонемы или графемы) [53].

Вторая группа уровней оперирует со сравнительной сложностью языковой единицы. Ее выделение основано на предположении, что единицы языка — фонемы (графемы) и морфемы — образуют не просто последовательные цепочки, а объединяются в интегральные единицы, имеющие более высокий порядок сложности и функционирующие как целостные элементы. Такие элементы характеризуются определенными всеобщими свойствами, превосходящими характеристики простой суммы составляющих единиц. П. Гарвин называет такие интегральные единицы «слитными единицами» (fused units). Уровни второй группы называются *уровнями интеграции*.

Третья, высшая группа уровней естественного языка оперирует со способом организации языковых единиц (элементарных и «слитных»). Любая знаковая система по определению содержит более одного знака. При любом использовании знаковой системы необходимо производить выбор знаков (selection) из числа имеющихся в этой системе. Этот принцип выбора применим и к естественному языку. Всякий раз, когда используется естественный язык, необходимо выбрать одну или несколько языковых единиц. Во многих знаковых системах выбор знаков является единственным организующим принципом. Примером такой знаковой системы могут служить обычные дорожные знаки. Однако в естественном языке функция каждой цепочки языковых единиц может быть тесно связана или даже определяться присутствием в этой цепочке других языковых единиц. Таким образом, нужно не только выбрать языковые единицы, но и расположить их в определенной последовательности. Хороший пример необходимости выполнения этого требования — английские предложения *dog bites man* [собака кусает человека] и *man bites dog* [человек кусает собаку]. В этих предложениях выбраны одни и те же языковые единицы Но из-за того что порядок расстановки этих единиц в предложениях различен, они имеют разный смысл. Поэтому в естественном языке применяются два принципа организации — выбор и расположение. Эти принципы — суть уровни организации естественного языка. Уровень выбора и уровень расположения можно также рассматривать соответственно как парадигматический и синтагматический уровни языка.

Итак, «язык — это знаковая система, структура которой специфицируется в терминах трех групп уровней, а именно — двух уровней построения, двух уровней организации и более чем одного уровня интеграции» [54]. По мнению П. Гарвина, это

[53] Графема — это минимальная структурная единица письменной речи, соответствующая фонеме в устной речи и служащая для различения слов.

[54] P. L. Garvin. Указ. соч., стр. 21.

определение может быть применено к любой знаковой системе с целью определения, является ли данная система естественным языком (или его полным изоморфом) или нет. Если знаковая система не имеет хотя бы одного из названных в определении свойств, то она не является естественным языком.

Полное описание ИПЯ на базе модели П. Гарвина или какой-либо другой лингвистической модели потребует его детального анализа с применением соответствующих лингвистических методов. Если удастся выполнить такой анализ, то станет возможной оценка структуры любого ИПЯ и сравнение разных ИПЯ на основе общих, четко определенных принципов. В связи с этим необходимо отметить, что существенное отличие любого ИПЯ от естественного языка состоит в том, что на графемном уровне (ИПЯ — письменные языки, не имеющие фонетики) в нем обычно используется смесь букв, цифр и специальных знаков, многие из которых не имеют звуковых обозначений. Хотя большинство ИПЯ не обладает «естественной» фонемной структурой, многие из них имеют остальные уровни построения, интеграции и организации (по Гарвину). Однако действие разных ИПЯ на этих уровнях носит разный характер, и это может оказывать сильное влияние на эффективность ИПС.

Попытки построить обобщенную теорию индексирования предпринимаются Ф. Йонкером (США) с 1957 г.[55] В этой теории Ф. Йонкер рассматривает информационный поиск как «поиск по совпадению терминов» и считает определяющими факторами любой ИПС тип используемых в ней лексических единиц и способы выражения логических связей между понятиями, обозначенными этими лексическими единицами [56]. Соответственно он постулирует существование двух континуумов (спектров) — терминологического и связанности. Если проинтегрировать эти два континуума, то любую ИПС можно представить на плоскости в виде точки с координатами ($x$, $y$), где $x$ — примененный в данной ИПС метод индексирования (от координатного до иерархической классификации), а $y$ — тип лексических единиц, используемых при таком методе индексирования.

*Терминологический континуум* (terminological continúúm) образован совокупностью лексических единиц ИПЯ или так называемых *термов* (terms), расположенных по степени возрастания их средней длины. Под термом Ф. Йонкер понимает лексическую единицу любого ИПЯ — слово, словосочетание, предметный заголовок или иерархическую цепочку слов, используемую для

[55] F. Jonker. The descriptive continuum: a «generalized» theory of indexing. Preprints of papers for the International conference on scientific information (Washington. D. C., November 16—21, 1958). Washington, National Academy of Sciences—National Research Council, Area 6, стр. 21—41.

[56] F. Jonker. Indexing theory, indexing methods and search devices. New York, The Scarecrow Press, 1964.

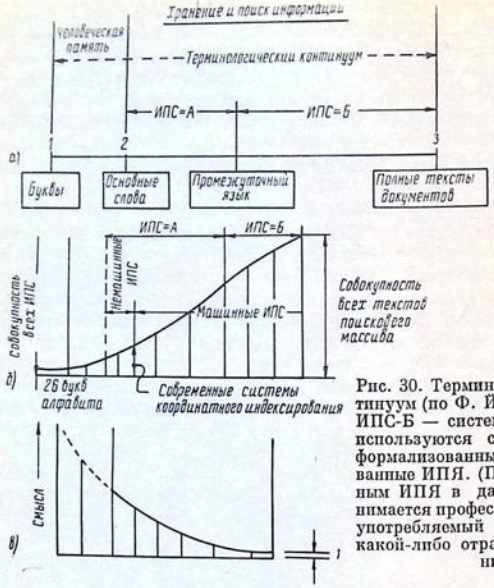



Рис. 30. Терминологический континуум (по Ф. Йонкеру) ИПС-А и ИПС-Б — системы, в которых используются соответственно неформализованные и формализованные ИПЯ. (Под формализованным ИПЯ в данном случае понимается профессиональный язык, употребляемый специалистами какой-либо отрасли науки и техники)

обозначения основного смыслового содержания документов. Длина терма определяется числом входящих в него слов. Минимальная длина терма равна букве того или иного алфавита, а максимальная — тексту целого документа.

Терминологический континуум схематично показан на рис. 30 *а*, *б*. Из этого рисунка видно, что крайнюю левую точку такого континуума образует небольшая совокупность букв алфавита какого-либо естественного языка, а крайнюю правую точку — потенциально бесконечное число термов (совокупность текстов всех документов, образующих поисковый массив). На рис. 30*в* показано изменение величины интенсионала (смысла) терма в зависимости от его длины. В крайней правой точке континуума интенсионал терма равен единице (здесь в качестве отдельных термов выступают тексты целых документов), а в крайней левой точке он стремится к бесконечности. Ф. Йонкер считает, что использование крайних точек терминологического континуума маловероятно и что в большинстве ИПС применяются языки, расположенные ближе к левому краю этого континуума.

*Континуум связанности* (connective continuum) схематично показан на рис. 31. Положение каждого ИПЯ в этом континууме (как и в терминологическом континууме) определяется средней длиной термов. Из определения терма можно сделать вывод, что его длина прямо пропорциональна степени иерархической связанности слов, образующих этот терм, и обратно пропорциональна степени его координативной связанности. По мнению

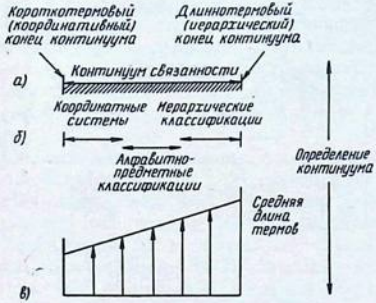



Рис. 31. Схема континуума связанности

Ф. Йонкера, в большинстве ИПЯ используются всего лишь два типа связи между лексическими единицами — иерархическая и/или координативная. Например, в предметных заголовках слова связаны иерархическими отношениями. Когда же документ индексируется несколькими предметными заголовками, то между ними существует координативная связь.

На рис. 31 *б* показано относительное положение трех основных типов ИПЯ — иерархических классификаций, алфавитно-предметных классификаций и языков дескрипторного типа — в континууме связанности. Крайнюю левую часть континуума образуют дескрипторные языки, в которых используются самые короткие термы. Этой части континуума соответствуют ИПЯ с посткоординацией. По мере перемещения по континууму вправо средняя длина терма увеличивается. Центральная часть континуума образована алфавитно-предметными классификациями, т. е. ИПЯ с частичной предкординацией, а ее крайняя правая часть — иерархическими классификациями (полная предкоординация). В иерархических классификациях средняя длина термов достигает максимума, так как в них каждый терм состоит из длинной цепочки иерархически связанных слов. При использовании иерархических классификаций поисковый образ документа может состоять всего лишь из одного терма.

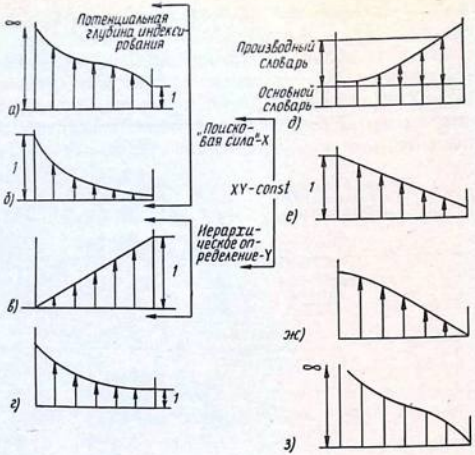



**Рис. 32. Изменение величины рабочих параметров континуума связанности в зависимости от средней длины терма**

Кроме главного или определяющего параметра — средней длины терма, континуум связанности обладает также следующими производными (рабочими) параметрами, которые суть функции главного параметра:

*а*) потенциальная глубина индексирования;

*б*) возможность индексирования по разным сочетаниям характеристик;

*в*) степень иерархического определения;

*г*) число поисков, необходимых для решения поисковой задачи;

*д*) объем словаря;

*е*) потребность в механизме координации термов;

*ж*) возможность ложных координаций двух и более термов;

*з*) разрешающая способность в отношении семантической неоднозначности.

Зависимость каждого из этих параметров от средней длины терма показана на рис. 32.

Самый важный из рабочих параметров — это потенциальная глубина индексирования. Величина этого параметра (см. рис. 32 *а*) пропорциональна числу характеристик, которые можно использовать при индексировании того или иного документа. Для идеальной линейной классификации этот параметр равен единице, так как при индексировании по такой классификации невоз-

можно применить больше характеристик, чем их было использовано при построении выбранного уровня данной классификации. По мере перемещения по континууму влево возрастает число характеристик, которые можно использовать при индексировании какого-либо документа. В крайней левой точке континуума величина этого параметра стремится к бесконечности.

Второй рабочий параметр континуума связанности — возможность индексирования по разным сочетаниям характеристик (рис. 32 *б*). При использовании иерархических классификаций поисковый образ документа не может содержать более одного терма. Число возможных перестановок термов в таком поисковом образе равно нулю. С уменьшением средней длины термов возможность индексирования по разным сочетаниям характеристик увеличивается; она становится полной в крайней левой точке континуума.

От этих двух параметров зависит, насколько точно может быть произведено индексирование документов и их последующий поиск. Вместе они определяют то, что Ф. Йонкер называет «поисковой силой» (retrieval power) системы индексирования.

Третий рабочий параметр — степень иерархического определения поисковых образов (рис. 32 *в*). Для линейных классификаций характерно 100-процентное иерархическое определение поисковых образов, так как индексирование по любой одной характеристике означает также индексирование по всем другим характеристикам, которые были использованы для образования вышестоящих уровней иерархии. В крайней левой точке континуума иерархическая связанность совсем исчезает из ИПЯ. Полное отсутствие иерархической связанности вызывает опасность невыдачи документа при поиске, если в поисковом предписании использован терм (ключевое слово или дескриптор), не относящийся к тому уровню иерархии, к которому принадлежит терм, примененный для индексирования этого документа.

При выборе термов для построения поискового образа того или иного документа возможны лишь две степени свободы: а) характеристика, по которой производится индексирование; б) уровень иерархии в пределах этой характеристики. Первая степень свободы, как это видно из рис. 32 *а* и 32 *б*, соответствует «поисковой силе» системы индексирования, а вторая степень — мере иерархического определения (рис. 32 *в*). Если обозначить «поисковую силу» системы индексирования через $X$, а степень иерархического определения поисковых образов через $Y$, то тогда связь между этими параметрами может быть приблизительно описана следующим образом:

$$X \cdot Y = \text{const}.$$

Четвертый рабочий параметр континуума связанности — это количество поисков, необходимое для решения каждой поиско-

вой задачи (рис. 32 *г*). Если в ИПС используется система индексирования, расположенная в левой части континуума, где степень иерархического определения поисковых образов невелика, то для осуществления исчерпывающего информационного поиска может потребоваться несколько подпоисков по термам, принадлежащим к разным уровням иерархии. Например, поиск по двум термам, относящимся к разным категориям и имеющим по несколько уровней иерархии, необходимо производить по всем возможным перестановкам характеристик. Наоборот, в крайней правой точке континуума для решения поисковой задачи достаточно произвести лишь один поиск.

Пятый рабочий параметр — объем словаря ИПЯ (рис. 32 *д*). Ф. Йонкер называет этот параметр «аппаратом доступа» (access apparatus). Левому краю континуума соответствуют ИПЯ, словари которых состоят из алфавитных списков ключевых слов. Объем таких словарей сравнительно невелик. В языках, соответствующих правой части континуума, словарный состав содержит не только отдельные слова, но и их различные сочетания. Очевидно, что количество таких сочетаний может быть очень большим.

Шестой рабочий параметр — необходимость в механизме координации термов (рис. 32 *е*). При использовании ИПЯ, находящегося у левого края континуума, обязательно нужен механизм координации термов. Если же применяемый ИПЯ находится у правого края континуума, то такой механизм не нужен. Для алфавитно-предметных классификаций (центральная часть континуума) механизм координации может потребоваться для логического умножения предметных заголовков.

Еще один рабочий параметр континуума связанности — возможность образования ложных координаций (рис. 32 *ж*). Величина этого параметра равна нулю в крайней правой точке континуума и максимальна в крайней левой точке. Образование ложных координаций приводит к выдаче документов, не отвечающих на информационный запрос, т. е. к появлению «поискового шума».

Последний из ранее названных рабочих параметров — разрешающая способность ИПЯ в отношении семантической неоднозначности. Из рис. 32 *з* видно, что чем меньше средняя длина терма, тем выше такая разрешающая способность. В крайней левой точке континуума величина этого параметра достигает почти бесконечности. При использовании ИПЯ, находящегося у левого края континуума, путем сочетания соответствующих термов можно выражать даже самые сложные понятия.

Ф. Йонкер подчеркивает, что приведенные им кривые носят лишь ориентировочный характер и требуют уточнения. Кроме рабочих параметров, он называет еще два «стоимостных» параметра континуума связанности, определяемые следующим образом: а) стоимость индексирования наиболее низка на левом краю

континуума связанности и наиболее высока — на его правом краю; б) стоимость информационного поиска наиболее низка на правом краю континуума (иерархические классификации) и наиболее высока на левом (короткотермовом) краю континуума.

Основным критерием для выбора рабочей точки в континууме связанности Ф. Йонкер считает степень диффузности, которая свойственна информации, содержащейся в индексируемых документах. Он пишет: «Вследствие диффузности информации ее документальные источники (items of information) могут представлять интерес для огромного множества проблем, применений и областей чистой и прикладной науки. Это означает, что при индексировании таких источников должна быть отражена каждая потенциально важная точка зрения. Поэтому диффузность какого-либо источника измеряется числом возможных точек зрения или критериев, используемых при индексировании этого источника. Высокая степень диффузности технологической информации не случайна. Она присуща самой природе нашей технической цивилизации, в которой каждое открытие или достижение в одной области может представлять интерес почти для любой другой области.

В большой диффузности информации состоит сущность проблемы индексирования и главный фактор в обобщенной теории индексирования. Если создать информационный массив (information collection) лишь для весьма специальных целей, то при его индексировании будет охвачена только часть всех потенциальных использований информации. Когда же указатель предназначается для любых возможных применений, то индексирование по своей глубине должно охватить все предвидимые ситуации» [57].

Ф. Йонкер считает, что если информация отличается небольшой диффузностью, не требующей применения более одного-двух разных критериев (аспектов) индексирования, и если эти критерии всегда имеют одно и то же иерархическое подчинение, то теоретически возможно работать в любой точке континуума и использовать либо библиотечно-библиографические классификации, либо дескрипторные языки. Когда же для информации характерна большая диффузность, то единственная возможность — использование ИПЯ, находящегося у левого края континуума связанности, так как только тогда будет возможно индексирование по любому сочетанию характеристик и на любую глубину.

В настоящее время еще трудно высказать какое-либо определенное суждение о теории Ф. Йонкера. В целом показанные на рис. 32 зависимости рабочих параметров континуума связанности как будто не противоречат практическому опыту. Но предстоит

[57] F. Jonker. Indexing theory, indexing methods and search devices. New York, The Scarecrow Press, 1964, стр. 52—53.

диться лишь на основе сравнения двух текстов — поискового образа документа и поискового предписания. При этом предполагается, что документы с идентичным содержанием заиндексированы так же идентично. Но такое требование в реальных условиях пока не выполнимо. Отсутствие единообразия в индексировании и его неадекватность — основные причины недостаточной эффективности документальных ИПС. Подчеркивая неадекватность любой системы индексирования, известный математик Й. Бар-Хиллел отмечал: «Для систем поиска литературы документу ставится в соответствие некоторое множество символов. Одна и притом единственная функция этих символов — служить лицу, производящему поиск литературы, ключом для оценки релевантности этого документа. Не будет неправильным сказать, что эти символы — будь то отдельные слова, целые фразы, предметные заголовки, дескрипторы, унитермы и т. д. — . . . как-то *обозначают* информационное содержание документа. Однако скорей в заблуждение вводят часто встречающиеся утверждения, что такие символы сами *содержат* — например, в конденсированном виде — часть информации, приведенной в этом документе» [63].

В связи с этим большое значение приобретает проблема формализации процедуры индексирования. Такая формализация не обеспечивает адекватности индексирования документов и информационных запросов. Но она обеспечивает единообразие индексирования, и уже это значительно повышает эффективность документальных ИПС.

Из признания субъективного характера оценок релевантности следует, что процедура ее определения не поддается полной формализации. Очевидно, что наилучшие результаты могут быть получены тогда, когда потребитель сам оценивает релевантность документов, выдаваемых ему той или иной системой. Для этого необходимо присутствие потребителя при проведении поиска документов по его информационному запросу и даже личное участие в поиске (например, с целью корректировки процесса поиска по его промежуточным результатам).

Однако такое требование невыполнимо в условиях работы большинства крупных документальных ИПС, которые создаются в информационных центрах и обычно рассчитаны на обслуживание большого числа потребителей. Многие потребители находятся в значительном удалении от информационного центра. Свои запросы они направляют в этот центр по почте и другим каналам связи. Индексирование таких запросов производится лишь по их текстам, причем потребителю не могут быть заданы вопросы, уточняющие характер нужной ему информации. Поэтому в крупных ИПС оценка релевантности документов должна осуществ-

[63] Y. Bar-Hillel. A logician's reaction to theorizing on information search systems. «American Documentation», 1957, v. 8, № 2, стр. 105.

ляться по формальным правилам, обеспечивающим необходимое единообразие в этих оценках, но не требующим непременного участия потребителя в проведении поиска документов по его информационному запросу.

Из сказанного выше можно сделать вывод, что для документальных ИПС, создаваемых в крупных информационных центрах, наиболее сильным критерием смыслового соответствия будет такой, который позволяет в максимальной степени учитывать специфику каждого информационного запроса и в то же время допускает формализацию процедуры определения релевантности документов. В связи с этим необходимо отметить, что формальная оценка величины релевантности документов и эшелонирование их выдачи практически осуществимы лишь тогда, когда ИПС создается на базе ЭЦВМ. Попытки решать эти задачи при помощи более простых технических средств обычно не приводят к удовлетворительным результатам.

Простейший и наиболее распространенный критерий смыслового соответствия состоит в требовании полного совпадения поискового образа документа с поисковым предписанием. Такой критерий применяется во всех документальных ИПС, основанных на использовании библиотечно-библиографических и алфавитно-предметных классификаций, а также в большинстве ИПС с дескрипторными языками. Однако при проведении информационного поиска с помощью систематического или алфавитно-предметного каталога (указателя) потребитель значительно усиливает этот критерий, так как в процессе поиска содержательно корректирует и уточняет его результаты по предметным подзаголовкам, предметным записям и/или библиографическим описаниям документов, обычно содержащимся в таком каталоге (указателе). Это — очень ценное свойство систематических и алфавитно-предметных каталогов (указателей), существенно повышающее эффективность производимого при их помощи поиска документов.

Возможность содержательной корректировки процесса поиска, как правило, отсутствует в документальных ИПС, основанных на использовании дескрипторных языков. В большинстве таких ИПС поисковые образы документов ставятся в соответствие не библиографическим описаниям или текстам этих документов, а лишь адресным шифрам их хранения. Поэтому для содержательного суждения о степени релевантности какого-либо документа необходимо извлечение его оригинала или копии из хранилища и предъявление этого документа потребителю. Это важное обстоятельство нельзя упускать из виду при любых сравнениях эффективности документальных ИПС разных типов.

В реальных ИПС вероятность полного совпадения поисковых образов документов с тем или иным поисковым предписанием очень низка. Поэтому необходимо иметь возможность как-то оценивать степень их частичного совпадения. Если такое совпа-

дение достигает необходимой и достаточной величины, то документ считается релевантным и выдается в ответ на информационный запрос. Рассмотрим несколько вариантов такого критерия смыслового соответствия для ИПС, основанных на использовании дескрипторных языков. Для целей данного изложения дескрипторный ИПЯ достаточно определить как язык, лексика которого состоит из определенного множества знаменательных слов естественного языка — *дескрипторов* (с устраненной синонимией, полисемией и омонимией этих слов), а синтаксические средства — из логических отношений И, ИЛИ и НЕ. Более точное определение дескрипторного ИПЯ дано в главе 6.

Обозначим через $R$ степень совпадения между поисковым образом документа и поисковым предписанием, необходимую и достаточную для того, чтобы считать этот документ в основном отвечающим на информационный запрос. Величину $R$ будем выражать в процентах. Кроме того, пусть $p$ — общее число дескрипторов в поисковом образе документа и $c$ — число дескрипторов поискового образа, совпадающих с дескрипторами поискового предписания. Тогда

$$R = \frac{c}{p} 100\%.$$

В простейшем случае, когда требуется полное совпадение поискового образа документа с поисковым предписанием ($c=p$), $R=100\%$. Однако в практических условиях достаточным может быть $R \geqslant 25\%$. Например, в системе избирательного распределения информации, которая создана и используется Центром технической информации Отдела систем обработки данных фирмы «Интернэшнал бизнесс машинз, корп.» (г. Покипси, штат Нью-Йорк), приемлемым считается $R$, величина которого может изменяться в пределах от 25 до 40%. Эта система реализована на ЭЦВМ типа ИБМ-7090. Допустим, что $p=12$ и $R=30\%$. Тогда ИПС должна выдавать документы, в поисковых образах которых содержится всего по четыре дескриптора, совпадающих с дескрипторами поискового предписания [64].

Возможны, по крайней мере, два простых способа усиления этого критерия смыслового соответствия. Первый способ состоит в том, что ИПС вместе с адресными шифрами документов, для которых выполнен формальный критерий смыслового соответствия, выдает также полные поисковые образы этих документов. По этим поисковым образам потребитель корректирует результаты работы ИПС. Такой способ применяется в ИПС «Уолнат», ко-

---

[64] R. J. Tritschler. A computer-integrated system for centralized information dissemination, storage, and retrieval. «Aslib Proceedings», 1962, v. 14, № 12, стр. 485.

торая с 1961 г. используется в Центральном разведывательном управлении США [65].

Второй способ усиления критерия смыслового соответствия, основанного на требовании частичного совпадения поискового образа документа с поисковым предписанием, состоит в введении так называемых «весовых» коэффициентов (weights). Сущность этого способа заключается в следующем. При формулировании поискового предписания потребитель оценивает каждый дескриптор этого предписания определенным «весовым» коэффициентом. Для этого он пользуется специальной шкалой, которая может содержать как положительные, так и отрицательные значения. Кроме того, потребитель может назвать численную величину $R$, но теперь она будет выражаться не в процентах, а в некоторых условных единицах (баллах). Документ считается релевантным, если сумма «весовых» коэффициентов тех дескрипторов поискового предписания, с которыми совпали дескрипторы поискового образа этого документа, не меньше установленной величины $R$.

Допустим, что в ИПС применяется 18-балльная шкала «весовых» коэффициентов, имеющая 9 положительных и 9 отрицательных делений. Поисковое предписание сформулировано в виде перечня дескрипторов $A$, $B$, $C$, $D$, $E$ и $F$; $A=6, B=9, C=1, D=3$; $E=9$ и $F=6$; $R \geqslant 15$. Тогда возможны, например, две такие ситуации:

а) в поисковом образе документа содержатся дескрипторы $A$, $B$ и $D$, для которых $R=12$. Это означает, что данный документ не отвечает на поставленный информационный запрос;

б) в поисковом образе документа содержатся дескрипторы $B$, $E$ и $F$, для которых $R=24$. Такой документ, безусловно, отвечает на информационный запрос и выдается системой [66].

Описанный выше критерий смыслового соответствия был применен для системы избирательного распределения информации, созданной в северо-западном отделении фирмы «Интернэшнал бизнесс машинз, корп.» (г. Чикаго) [67]. В этой системе, которая реализована при помощи ЭЦВМ типа ИБМ-1401, средний объем поискового образа одного документа составляет 22 или 41 дескриптор (в зависимости от метода индексирования), а средний

---

[65] P. D. Bradshaw. The Walnut system: a large capacity document storage and retrieval system. «American Documentation», 1962, v. 13, № 3, стр. 274.

[66] R. J. Tritschler. A computer-integrated system for centralized information dissemination, storage, and retrieval. «Aslib Proceedings», 1962, v. 14, № 12, стр. 498.

[67] K. D. Ofer. Selective information dissemination and retrieval at Chemicals and Phosphates Ltd. with an IBM-1401 DPS. В кн.: «Congresso internazionale sulla documentazione e informazione scientifico-technica». Roma, 2—11 febbraio 1964, v. 1.

Примерно такой же критерий смыслового соответствия применен для системы избирательного распределения информации, созданной в Бюро мелиорации Министерства внутренних дел США (1963 г.). В этой системе каждый документ обычно индексируется 15—25 дескрипторами. Звездочкой в поисковом образе документа отмечаются те дескрипторы, которые наиболее точно выражают основное смысловое содержание этого документа (в среднем 3,2 дескриптора, но не более 4). Кроме того, поисковый образ каждого документа может быть дополнен так называемыми «идентификаторами» (identifiers) — именами собственными, фирменными и географическими названиями и т. д., не являющимися дескрипторами.

В поисковом предписании, которое не должно содержать более 18 дескрипторов (его средний объем — 16,9 дескрипторов), потребитель также отмечает звездочками не более 4 дескрипторов, несущих в его информационном запросе наибольшую смысловую нагрузку. При совпадении хотя бы одного дескриптора поискового образа документа с дескриптором поискового предписания возможны такие ситуации:

1. Неотмеченный дескриптор поискового образа совпал с неотмеченным дескриптором поискового предписания.
2. Отмеченный дескриптор поискового образа совпал с неотмеченным дескриптором поискового предписания.
3. Неотмеченный дескриптор поискового образа совпал с отмеченным дескриптором поискового предписания.
4. Отмеченный дескриптор поискового образа совпал с отмеченным дескриптором поискового предписания.

Если обозначить вероятные степени релевантности документа для случаев (1)—(4) соответственно символами $R_1$—$R_4$, то $R_4 > R_3 > R_2 > R_1$. Было решено считать документ релевантным, если хотя бы один дескриптор поискового образа данного документа совпадает с дескриптором поискового предписания и если хотя бы один из этих дескрипторов отмечен [70]. Релевантным считается также документ, в поисковом образе которого содержится не менее 3 любых дескрипторов, совпадающих с дескрипторами поискового предписания (отмеченность в данном случае никак не учитывается)[71].

Метод выделения наиболее важных дескрипторов в поисковых образах документов и поисковых предписаниях применяется

[70] J. W. Hill. Matching of descriptors in a selective dissemination system. В кн.: «Automation and scientific communication». Pt. 1. Washington, American Documentation Institute, 1963, стр. 65—66.

[71] Thesaurus of descriptors. A list of keywords and cross-references for indexing and retrieving the literature of water resources development. Tentative edition. Denver, Colo., U. S. Department of Interior, Bureau of Reclamation, 1963, стр. 132—133.

также в Военном документационном центре США и в Федеральном центре по распространению научной и технической информации США. В частности, отмеченные дескрипторы поисковых образов научно-технических отчетов, аннотируемых в бюллетенях «Technical Abstract Bulletin» и «U. S. Government Research and Development Reports», которые выпускаются этими информационными центрами, используются в качестве предметных заголовков для алфавитно-предметных указателей к названным бюллетеням.

В документальной ИПС, которая создана в Центре научной и технической информации Национального управления по аэронавтике и исследованию космического пространства США (г. Бесезда, штат Мэриленд), метод «весовых» коэффициентов применяется не только для усиления критерия смыслового соответствия и эшелонирования выдачи. Оказалось, что этот метод позволяет моделировать решения соответствующих уравнений булевой алгебры. Если правильно подобрать численные значения «весовых» коэффициентов для дескрипторов и $R$ поискового предписания, то можно получить точно такой же результат, какой дает нам решение соответствующего уравнения булевой алгебры. Например, уравнение $A \,\&\, (B+C+D)=$ [ответ] можно заменить подстановкой следующих «весовых» коэффициентов: $A=3$, $B=1$, $C=1$, $D=1$; $R=4$. Символы & и + здесь и далее употребляются для обозначения соответственно логических связок И и ИЛИ.

Использование этого важного свойства метода «весовых» коэффициентов позволяет получать существенную экономию машинного времени при информационном поиске. Как показывает опыт, для вычисления $R$ по «весовым» коэффициентам дескрипторов поискового предписания обычно требуется в четыре раза меньше машинного времени, чем для решения соответствующего уравнения булевой алгебры (ИПС реализована на ЭЦВМ типа ИБМ-1410). Основные типы булевых формул, которые допускают такое моделирование, и необходимые для этого численные значения «весовых» коэффициентов, приведены в табл. 17.

Была разработана также система простейших правил для определения того, какие численные значения должны иметь «весовые» коэффициенты и $R$, когда поисковое предписание имеет более сложную логическую структуру, чем приведенные в табл. 17 типовые формулы.

Однако имеется ряд булевых формул, которые невозможно моделировать описанным выше способом, т. е. путем приписывания дескрипторам положительных или отрицательных «весовых» коэффициентов. Основными из таких формул являются:

$$(A + B)\,\&\,(C + D). \tag{7}$$

$$(A\,\&\,B) + (C\,\&\,D). \tag{8}$$

Таблица 17

**Численное значение «весовых» коэффициентов для моделирования решения булевых формул**

| Порядковый номер | Булевы формулы | «Весовые» коэффициенты | | | | $R$ |
|---|---|---|---|---|---|---|
| | | $A$ | $B$ | $C$ | $D$ | |
| 1 | $A \& (B+C+D)$ | 3 | 1 | 1 | 1 | 4 |
| 2 | $A+B+C+D$ | 1 | 1 | 1 | 1 | 1 |
| 3 | $A \& B \& C \& D$ | 1 | 1 | 1 | 1 | 4 |
| 4 | $A+(B \& C \& D)$ | 3 | 1 | 1 | 1 | 3 |
| 5 | $(A+B)+(C \& D)$ | 2 | 2 | 1 | 1 | 2 |
| 6 | $(A+B) \& (C \& D)$ | 1 | 1 | 2 | 2 | 5 |

Для таких формул был разработан метод моделирования по «весовым» коэффициентам, которые приписываются группам переменных. Сущность этого метода состоит в развертывании формулы, выделении в ней определенных структурных групп и приписывании членам каждой группы соответствующих «весовых» коэффициентов и $R$. Например, формула (7) в результате такого преобразования будет иметь следующий вид:

$$A \& (C + D) + B \& (C + D),$$

где $A \& (C+D)$ — группа $A$ $(A=3;\ C=1;\ D=1;\ R=4)$ и
$B \& (C+D)$ — группа $B$ $(B=3,\ C=1,\ D=1;\ R=4)$.

Более сложные уравнения булевой алгебры выгодней решать обычным путем, так как процедура их кодирования становится слишком трудной для человека. В связи с этим необходимо отметить, что в данной ИПС формулирование поисковых предписаний в виде уравнений булевой алгебры не является обязательным. Метод «весовых» коэффициентов в ней можно применять и в чистом виде [72].

Сильный критерий смыслового соответствия применен в документальной ИПС, которая в 1961 г. была создана в фирме «Сен-Гобэн» (Франция)[73]. ИПС реализована при помощи ЭЦВМ типа ИБМ-705. В ней потребитель не только задает путем перечисления множество дескрипторов, выражающее основное смысловое содержание его информационного запроса, но также соответствующим образом упорядочивает это множество, располагая

[72] W. T. Brandhorst. Simulation of Boolean logic constraints through the use of term weights. «American Documentation», 1966, v. 17, № 3, стр. 145—146.

[73] F. Levery, B. Reinbell, R. Cohendet et al. Une expérience de sélection automatique de documentation. «Economie et comptabilité gestion des enterprises», 1961, v. 15, № 54, стр. 2641—2663.

названные дескрипторы в порядке убывания их смысловой нагрузки в информационном запросе. Это эквивалентно приписыванию дескрипторам поискового предписания соответствующих «весовых» коэффициентов [74]. Кроме того, потребитель устанавливает разные уровни точности выдачи (т. е. определяет ее эшелоны) и таким образом сам контролирует точность и полноту информационного поиска, проводимого по его запросу.

Сущность применяемого в ИПС фирмы «Сен-Гобэн» критерия смыслового соответствия можно раскрыть следующим образом. Пусть $A$, $B$, $C$, $D$, $E$, $F$ и т. д. — дескрипторы поискового предписания, расположенные потребителем в порядке убывания их важности. Тогда выдачу первого уровня точности образуют документы, в поисковые образы которых входят дескрипторы $A$ или $A \& B$ или $A \& B \& C$ и т. д., т. е. $m$ первых дескрипторов поискового предписания. Очевидно, что в этой выдаче содержится максимальное число нерелевантных документов, но также и максимум релевантных документов (наибольший объем и широта выдачи).

Выдачу второго уровня точности образуют документы, в поисковые образы которых входят $n$ первых дескрипторов поискового предписания ($n > m$). В этой выдаче процентное содержание нерелевантных документов будет значительно ниже, чем в выдаче первого уровня точности. В выдаче следующего уровня точности «поисковый шум» будет еще ниже и т. д. Однако чрезмерное повышение точности может привести к потере релевантных документов, т. е. к сокращению полноты выдачи. Поэтому в ИПС фирмы «Сен-Гобэн» применяется не более четырех уровней точности.

Необходимо обратить внимание на то, что описанный способ формулирования поисковых предписаний и критерия смыслового соответствия выполняет также определенные синтаксические функции. Это можно показать на следующем примере. Пусть даны два информационных запроса:

1. «Производство фталевой кислоты путем окисления ксилола с использованием бромистого магния в качестве катализатора».

2. «Использование бромистого магния в качестве катализатора при производстве фталевой кислоты путем окисления ксилола».

Поисковые предписания для этих обоих информационных запросов состоят из одних и тех же дескрипторов, хотя данные запросы по смыслу неэквивалентны один другому. Поэтому для обеспечения достаточно высокой точности выдачи по каждому из сформулированных выше запросов необходимо соответствую-

[74] General information manual. An information retrieval experiment by Saint-Gobain Company. Paris, France; White Plains, N. Y., IBM Data Processing Division, 1962.

щее уточнение поисковых предписаний. В ИПС фирмы «Сен-Гобэн» эти поисковые предписания будут иметь такой вид:

1. Фталевая кислота
   & Производство
   & Окисление
   & Ксилол — *Первый уровень точности*
   & Катализатор
   & Бромистый магний — *Второй уровень точности*

2. Бромистый калий
   & Катализатор — *Первый уровень точности*
   & Фталевая кислота
   & Производство
   & Окисление
   & Ксилол — *Второй уровень точности* [75]

Создатели ИПС «Пусто-Непусто-4» (ВИНИТИ) при выборе критерия смыслового соответствия для своей системы исходили из более общих представлений. Они поставили перед собой задачу сформулировать такой критерий смыслового соответствия, который бы обеспечивал достаточно эффективную работу их автоматической ИПС при любых информационных запросах. Для этого критерий смыслового соответствия должен быть полностью формальным, т. е. не должен требовать непременного участия потребителей в составлении поисковых предписаний или в оценке релевантности документов, выдаваемых им системой.

В документальной ИПС «Пусто-Непусто-4» критерий смыслового соответствия сформулирован в терминах пустоты и непустоты четырех описываемых ниже множеств $M_1$—$M_4$ (отсюда происходит и название этой системы):

$M_1$ — множество дескрипторов поискового образа, совпадающих хотя бы с одним из дескрипторов поискового предписания;

$M_2$ — множество дескрипторов поискового образа, в логическом смысле стоящих выше (но не обязательно непосредственно выше) хотя бы одного из дескрипторов поискового предписания;

$M_3$ — множество дескрипторов поискового образа, в логическом смысле стоящих ниже (но не обязательно непосредственно ниже) хотя бы одного из дескрипторов поискового предписания;

$M_4$ — множество дескрипторов поискового предписания, несравнимых ни с одним из дескрипторов поискового образа (в смысле отношения, упорядочивающего это множество дескрипторов).

Каждому из множеств $M_1$—$M_4$ поставлено в соответствие некоторое число $m_i$, которое определяется следующим образом:

$$m_i = \begin{cases} 0, \text{ если } M_i \text{ — пусто} \\ 1, \text{ если } M_i \text{ — непусто } (i = 1, 2, 3, 4). \end{cases}$$

[75] General information manual. An information retrieval experiment by Saint-Gobain Company. Paris, France; White Plains, N. Y. IBM Data Processing Division, 1962, стр. 11—12.

Тогда для любой пары «поисковый образ—поисковое предписание» можно составить 16 четырехразрядных двоичных чисел, каждое из которых будет характеризовать степень смыслового соответствия между документом и информационным запросом. Из этих 16 двоичных чисел (они суть номера соответствующих классов документов) были выбраны числа 0010, 0100, 0110, 1000, 1010, 1100 и 1110, которые обозначают классы, предположительно содержащие больше релевантных, чем нерелевантных документов. Далее эти числа были сгруппированы так, как показано в табл. 18.

Таблица 18

**Условия выдачи документов в ИПС «Пусто-Непусто-4»**

| Номер эшелона выдачи | Комбинации множеств $M_1 - M_4$ для документов, включаемых в эшелоны выдачи | | | | Номер эшелона выдачи | Комбинации множеств $M_1 - M_4$ для документов, включаемых в эшелоны выдачи | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $m_1$ | $m_2$ | $m_3$ | $m_4$ | | $m_1$ | $m_2$ | $m_3$ | $m_4$ |
| | 1 | 0 | 0 | 0 | 3 | 1 | 1 | 0 | 0 |
| 1 | 0 | 0 | 1 | 0 | | 0 | 1 | 1 | 0 |
| | 1 | 0 | 1 | 0 | 4 | 0 | 1 | 0 | 0 |
| 2 | 1 | 1 | 1 | 0 | | | | | |

В этой таблице эшелоны пронумерованы таким образом, чтобы вероятность выдачи искомых документов была максимальной в первом эшелоне и минимальной — в последнем, а вероятность выдачи нерелевантных документов («поисковый шум») — минимальной в первом эшелоне и максимальной в последнем [76]. Потребитель сам определяет, какие эшелоны должна выдать ИПС в ответ на его информационный запрос. Если потребителю не нужны все релевантные документы поискового массива, то он может ограничиться лишь первым эшелоном выдачи. Когда же требуется исчерпывающий поиск документов, то потребителю должны быть выданы все эшелоны [77].

Приведенная выше таблица ясно показывает, что в системе «Пусто-Непусто-4» условия выдачи второго и третьего эшелонов полностью совпадают. Поэтому эти два эшелона следовало бы объединить в один и говорить лишь о трех эшелонах выдачи. На это обстоятельство первым обратил внимание В. А. Успенский (1965 г.).

[76] Э. С. Бернштейн. Формализованный язык и критерий смыслового соответствия информационно-поисковой системы «Пусто-Непусто-4». Сб.: «Научно-техн. информация», 1963, № 12, стр. 37—39.

[77] Э. Бернштейн, Д. Лахути, В. Чернявский. Вопросы теории поисковых систем. М., ВНИИЭМ, 1966, стр. 22—23.

Большой интерес представляет критерий смыслового соответствия, примененный Г. Э. Стайлзом в документальной ИПС, которая была создана в Управлении национальной безопасности Министерства обороны США еще в 1956 г.[78] ИПС реализована на ЭЦВМ типа ИБМ-705. Для индексирования документов в ней используется система унитермов. Объем поискового массива — более 100 тыс. документов [79].

В этой ИПС для оценки релевантности документов применена алгоритмическая процедура, состоящая из следующих пяти шагов:

1. Сначала для каждого термина $t_i$ информационного запроса $Q_i$ составляется список терминов, которые совместно встречаются (ассоциируются) с $t_i$ в поисковых образах $\{p_d\}$ документов, введенных в данную ИПС. $Q_i=\{t_i\}$.

Использование в качестве меры ассоциативной связанности любых двух терминов непосредственно показателя частоты их совместной встречаемости дает неудовлетворительные результаты, так как величина такой частоты может в значительной мере зависеть от общей частоты употребления каждого из данных двух терминов в тексте по определенной тематике. Поэтому было решено применить для этой цели показатель относительной частоты, который обозначает величину отклонения от ожидаемой частоты совместной встречаемости любых двух терминов, вычисленной на основе предположения, что между этими терминами не существует никакой ассоциативной связи.

Для вычисления относительной частоты совместной встречаемости терминов из $\{p_d\}$ с $t_i \in Q_i$ (условимся обозначать ее символом $F_a$) была разработана следующая формула:

$$F_a = \log_{10} \frac{\left[(fN - AB) - \frac{N}{2}\right]^2 N}{AB\,(N-A)\,(N-B)},$$

где $A$ — число документов, в поисковые образы которых входит первый термин данной пары, но не входит второй термин; $B$ — число документов, в поисковые образы которых входит второй термин данной пары, но не входит первый термин; $f$ — число документов, в поисковые образы которых входят оба термина данной пары; $N$ — общее число документов в поисковом массиве.

Для каждого $t_i \in Q_i$ составляется список терминов, ассоциативно связанных с $t_i$ в $\{p_d\}$, причем в списке эти термины

---

[78] J. A. Sanford and F. R. Theriault. Problems in the application of uniterm coordinate indexing. «College and Research Libraries», 1956, v. 17, № 1, стр. 19—23.

[79] H. E. Stiles. The association factor in information retrieval. «Journal of the Association for Computing Machinery», 1961, v. 8, № 2, стр. 271—279.

располагаются в порядке убывания величины их $F_a$. Термины, для которых $F_a < 1$, в список не включаются. Такие списки, названные «профилями терминов» (term profiles), обладают следующими важными характеристиками:

а) «профили терминов» построены по фактическим поисковым образам документов, а не на основании каких-либо субъективных суждений. Поэтому в них входят лишь термины, полезные для информационного поиска;

б) «профили терминов» порождены алгоритмически;

в) «профили терминов» выявляют разные значения, в которых тот или иной индексирующий термин употребляется в данном собрании документов;

г) в «профили терминов» входят слова, связанные с $t_i$ не семантически, а лишь статистически. Статистическая связь между терминами обусловлена наличием определенного родства между денотатами, обозначенными этими терминами, а семантическая связь — наличием такого родства между концептами.

2. Следующий шаг процедуры — сравнение «профилей» всех терминов $Q_i$ и исключение из этих «профилей» таких терминов, которые одновременно не входят во все «профили» (или, по крайней мере, в некоторое заданное число «профилей»). Оставшиеся после такой проверки термины были названы «терминами первого поколения». Очевидно, что для $Q_i$, состоящего лишь из одного $t_i$, список терминов первого поколения совпадает с «профилем» $t_i$.

В числе терминов первого поколения обычно не содержатся синонимы и полусинонимы, так как в поисковый образ документа не могут одновременно входить даже два синонима одного и того же термина. Но синонимы и полусинонимы, а также родовые и видовые термины очень полезны для информационного поиска. Поэтому они должны быть добавлены к списку терминов первого поколения.

3. С этой целью шаги (1) и (2) повторяются для терминов первого поколения, которые в данном случае обрабатываются как термины информационного запроса. Число терминов первого поколения может быть довольно большим. Поэтому для отнесения ассоциирующихся с ним терминов к категории «терминов второго поколения» вполне достаточно одновременное их вхождение, например, лишь в $^1/_5$ всех «профилей». В число терминов второго поколения уже будут входить такие, которые связаны между собой не только статистически, но и семантически.

В результате выполнения описанных выше операций мы получим список, содержащий $\{t_i\}=Q_i$, а также все термины первого и второго поколения. Очевидно, что термины этого расширенного списка обладают разной степенью ассоциативной связанности с терминами $Q_i$. Поэтому для каждого $t_i \in Q_i$ полезно вычислить степень такой связи.

4. Для каждого термина расширенного списка вычисляется $F_a$ по отношению к другим терминам этого списка. Учитываются лишь значения $F_a > 1$. Сумма значений $F_a$ для каждого термина расширенного списка, деленная на число таких терминов в данном списке, определяет «вес» этого термина. Это позволяет расположить термины расширенного списка в порядке убывания их вероятной релевантности к $Q_i$.

5. Следующий шаг — сравнение терминов расширенного списка с терминами поискового образа каждого документа. При совпадении термина поискового образа какого-либо документа с термином расширенного списка этому документу приписывается «вес», который имеет данный термин в расширенном списке. Сумма таких «весов» для каждого документа образует «показатель релевантности документа», обозначающий степень смыслового соответствия между информационным запросом и содержанием документа. Выдаваемые документы располагаются в порядке убывания их показателей релевантности. Необходимо отметить, что описанная выше процедура позволяет выявлять в поисковом массиве такие релевантные документы, в поисковые образы которых не входят термины, образующие информационный запрос.

Для проверки эффективности данной процедуры был проведен такой эксперимент. Квалифицированному инженеру, который ничего не знал о проводившемся исследовании и его результатах, было поручено произвести содержательную оценку степени релевантности 56 документов по отношению к определенному информационному запросу. Степень совпадения его оценок с показателями релевантности документов была поразительно высокой.

Описанная Г. Э. Стайлзом алгоритмическая процедура определения релевантности документов весьма удобна при использовании стохастического метода индексирования, разработанного Г. П. Луном в 1957 г. Кроме того, эта процедура позволяет выявлять в потоке вводимых в ИПС документов такие, которые по содержанию значительно отличаются от документов, уже введенных в поисковый массив. Это может быть сделано путем сравнения «профилей» терминов информационного запроса, построенных по поисковым образам определенного множества новых документов, с «профилями», которые соответствуют уже входящим в поисковый массив документам. Если в первой группе «профилей» имеются термины, показатели ассоциативной связи которых с терминами $Q_i$ высоки, а во второй группе «профилей» такие показатели для данных терминов значительно ниже, то это свидетельствует об определенной новизне поступившего материала.

Для усиления критерия смыслового соответствия некоторые исследователи вводят в поисковые образы документов определенные формальные характеристики этих документов. В этом отношении весьма интересна работа М. Кесслера, который создал в Массачусетском технологическом институте (г. Кембридж,

штат Массачусетс) экспериментальную ИПС (реализована на ЭЦВМ типа ИБМ-7090). В эту ИПС уже введено свыше 40 тыс. журнальных статей по физике. Поисковый образ каждой статьи состоит из следующих характеристик: выходные данные статьи, ее заглавие; фамилии авторов статьи; название учреждения, в котором выполнена описываемая в статье работа; номер реферата или аннотации этой статьи в РЖ «Physics Abstracts»; предметные заголовки для этой статьи (если они даны в журнале).

Информационный поиск возможен по любой из перечисленных характеристик, а также по их разным сочетаниям. Для оценки релевантности документов весьма полезным оказалось введенное М. Кесслером понятие библиографического сочетания (bibliographic coupling). Как показывают проведенные им исследования, два документа можно считать по смыслу связанными между собой, если их авторы ссылаются на один или несколько одинаковых документов. Степень такой связанности определяется числом библиографических ссылок, совпадающих в этих двух документах [80].

Таким образом, возможны, например, такие формулировки информационных запросов [81]:

а) все статьи, аналогичные статье $A_i$;

б) все статьи, в заглавиях которых содержится слово $w_j$ и авторы которых ссылаются в основном на те же публикации, что и авторы статьи $A_i$.

Формальные характеристики документов (главным образом — имена авторов документов и содержащиеся в этих документах ссылки на другие документы) даже полностью заменили собой ИПЯ в системе сигнальной информации, которая создана фирмой «Институт фор сайентифик информейшн» (г. Филадельфия, США). Эта система, получившая название ASCA (Automatic Subject Citation Alert=Автоматическая система сигнальной информации на основе библиографических ссылок), начала функционировать с января 1965 г. В систему ASCA информационные запросы вводятся в виде таких характеристик, как фамилии авторов публикаций, интересующих потребителя; списки публикаций, ссылки на которые (одна или несколько) должны содержаться в релевантных публикациях; названия организаций и учреждений, в которых выполняются интересующие потребителя работы; имена владельцев патентов; классы и подклассы патентов.

В заключение данного раздела необходимо еще раз подчеркнуть, что определение понятия релевантности и тем более измерение ее величины — очень трудная задача. Оценивая

---

[80] M. M. Kessler. Bibliographic coupling between scientific papers. «American Documentation», 1963, v. 14, № 1, стр. 10—25.

[81] M. M. Kessler. The MIT technical information project. «Physics Today», 1965, v. 18, № 3, стр. 28—36.

перспективы решения этой задачи, Й. Бар-Хиллел писал, что определение понятия смысловой связанности документов в поисковом массиве, которая бы имела подходящую метрику для оценки степени такой связанности, «либо вообще недостижимо, либо в лучшем случае представляет собой дьявольски трудную задачу, причем ни один из известных подходов не имеет шансов привести к определению функции, которая даст нам полезную меру для оценки смыслового расстояния между темами, и еще меньше шансов — для оценки такого расстояния между документами» [82]. И далее: «. . . Я хотел бы в данном случае подчеркнуть, насколько трудно, если не просто невозможно, прийти к адекватному раскрытию *количественного аспекта* степени релевантности документа по отношению к теме. Это несомненно более трудная задача, чем раскрытие качественного аспекта этого понятия. Между прочим, понятие степени релевантности следует отличать от *вероятности быть релевантным*, хотя интуиция подсказывает, что между ними могут существовать какие-то плохо определенные связи» [83].

С такой крайне пессимистической точкой зрения нельзя согласиться. Для Й. Бар-Хиллела вообще характерна глубокая, но всегда только негативная критика любых попыток решить проблемы машинного перевода и автоматического информационного поиска. Вместе с тем необходимо определенно высказаться и против другой крайности — попыток найти решение этой труднейшей задачи в выявлении структурных корреляций между текстами поисковых образов документов и поисковых предписаний без учета их смысла и индивидуальной специфики каждого информационного запроса.

## МЕТОДЫ ОЦЕНКИ ЭФФЕКТИВНОСТИ ДОКУМЕНТАЛЬНЫХ ИПС

Рабочая характеристика любой системы (обозначим эту характеристику через $P$) измеряется в терминах функции двух переменных:

$$P = f(u, v),$$

где $u$ — техническая эффективность (effectiveness) и
$v$ — экономическая эффективность (efficiency).

Техническая эффективность определяется как мера способности системы выполнять те функции, для которых она предназначена, а экономическая эффективность — как мера стоимости

[82] Y. Bar-Hillel. Some theoretical aspects of mechanization of literature searching. Technical report No. 3. Jerusalem, Hebrew University, 1960, стр. 42.
[83] Там же, стр. 44.

выполнения этих функций. В данном разделе рассматривается главным образом проблема оценки технической эффективности документальных ИПС. Что же касается их экономической эффективности, то она измеряется какой-то стоимостной функцией $C$ от времени $t$:

$$v = C(t).$$

Таким образом $P$ может быть выражена в терминах технической эффективности $u$, приходящейся на единицу стоимости [84].

Будем далее считать, что функция документальной ИПС состоит в том, чтобы в ответ на информационный запрос (а не информационную потребность!) выделять из некоторого множества документов (поискового массива) такие, которые соответствуют этому запросу. Принадлежность документа к классу документов, подлежащих выдаче в ответ на данный информационный запрос, определяется на основании его релевантности.

Для количественной оценки эффективности любой системы необходимо определить эффективность ее идеальной модели, чтобы затем использовать величину идеальной эффективности в качестве меры для оценки эффективности реальных систем такого типа. По определению К. Муэрса (1959 г.), идеальной можно назвать ИПС, «выдающую каждому потребителю в ответ на его информационный запрос точно такое же множество документов (не больше и не меньше), какое отобрал бы сам этот потребитель, если бы он мог внимательно прочитать всю библиотеку и выбрать в ней документы, по его мнению, соответствующие информационному запросу» [85].

Моделирование работы идеальной ИПС обычно производится следующим образом. Создается группа судей, которые просматривают тексты всех документов поискового массива и совместно решают, какие документы система должна была бы выдать в ответ на предъявленный ей информационный запрос. Свое решение эти судьи принимают, исходя из текста информационного запроса. Но если формулировка информационного запроса — лишь приблизительное выражение информационной потребности ученого или инженера, то тогда подмножество документов, объявленных судьями релевантными по отношению к данному запросу, как правило, будет отличаться от того подмножества, которое бы отобрал потребитель, если бы он сам изучал каждый документ

---

[84] W. Goffman and V. A. Newill. Methodology for test and evaluation of information retrieval systems. CSL : TR-2. Cleveland, Ohio; Western Reserve University, Center for Documentation and Communication Research, 1964, стр. 6.

[85] C. N. Mooers. The intensive sample test for the objective evaluation of the performance of information retrieval systems. RADC-TN-59-160 (ZTB-132), Cambridge, Mass., Zator Co., 1959, стр. 5.

поискового массива. Р. Тейлор писал по этому поводу: «Например, несколько потребителей могут задать один и тот же вопрос — или предположительно один и тот же вопрос — и в ответ каждый получит одно и то же множество сообщений. Однако если каждый потребитель сам просмотрит весь поисковый массив, то он отберет из этого массива отличающееся от других подмножество ($S_1$, $S_2$. . . $S_k$) как релевантное по отношению к своему вопросу. Очевидно, что хотя их вопросы имели одно и то же словесное выражение, у каждого потребителя, по-видимому, была разная потребность. А мы предписываем каждому одно и то же лекарство» [86]. Таким образом, возникает вопрос об адекватности описанного выше метода моделирования работы идеальной ИПС, а вместе с ним — и вопрос об адекватности основанных на этом методе количественных оценок технической эффективности ИПС.

Итак, основой для оценки эффективности документальных ИПС является понятие релевантности. Но релевантность, как было показано в предыдущем разделе, — понятие субъективное. Оно представляет собой, в сущности, психологическую реакцию потребителя на конкретный результат работы ИПС. Опыт показывает, что в зависимости от условий места и времени даже один и тот же потребитель по-разному оценивает релевантность документов, выданных одной и той же системой в ответ на один и тот же информационный запрос. Слишком велики разногласия в оценках релевантности и тогда, когда применяется метод судейства. Поэтому возникает еще один вопрос, правомерно ли использование явно субъективного понятия релевантности в качестве основы для оценки эффективности документальных ИПС.

М. Таубе решительно возражал против незаконной подмены явно субъективного понятия релевантности математическим понятием релевантности как свойства ИПС. Он писал, что все полученные в результате такой подмены уравнения, кривые и математические формулы, которые якобы позволяют производить количественную оценку эффективности документальных ИПС, есть ни что иное, как псевдоматематика.

В письме к редактору журнала «American Documentation» по поводу статьи М. Таубе «О псевдоматематике релевантности», ранее опубликованной в том же журнале, Дж. Солтон оспаривал вывод М. Таубе о неправомерности использования релевантности в качестве основного критерия оценки эффективности ИПС [87]. Свои возражения Дж. Солтон основывает на том, что термин «релевантность» употребляется в двух разных значениях:

---

[86] R. S. Taylor. The process of asking questions. «American Documentation», 1962, v. 13, № 4, стр. 392.

[87] M. Taube. A note on the pseudo-mathematics of relevance. «American Documentation», 1965, v. 16, № 2, стр. 69—72; J. Salton. «American Documentation», 1965, v. 16, № 4, стр. 341.

а) *релевантность* 1 — это суждение потребителя о том, соответствует выданный поисковой системой документ его информационной потребности или нет;

б) *релевантность* 2 — это мера для оценки эффективности ИПС, определяемая как относительное число релевантных документов в выдаче. Дж. Солтон предлагает пользоваться для этой цели термином «точность» (precision).

Далее Дж. Солтон пишет: «Совершенно правильно, что точность (релевантность 2) зависит от релевантности 1, так как невозможно определить число релевантных документов в выдаче, если до этого кем-то не было установлено, какие документы релевантны по отношению к данному запросу и какие нерелевантны. Однако в равной степени верно и то, что когда высказано суждение о такой релевантности, скажем, некоторым данным потребителем $A$ (т. е. когда установлена релевантность 1), тогда точность (релевантность 2) есть совершенно однозначная мера, которую можно применять для определения эффективности системы с точки зрения потребителя $A$. Более того, если принять потребителя $A$ за типичного потребителя, то тогда измерение точности становится действительно ценным показателем эффективности системы также и для других потребителей.

С другой стороны, если потребитель $A$ не принимается в качестве типичного, то и тогда точность все же может быть определена однозначно. Это можно сделать путем оценки релевантности контрольной (test) группой потребителей и отметки данного документа как релевантного по отношению к определенному запросу, если с этим соглашается большинство членов этой группы. Эти суждения о релевантности путем согласования можно затем использовать, как и ранее, для вычисления величины полезной точности поиска».

Из приведенной выше цитаты ясно видна полная несостоятельность возражений Дж. Солтона. В этой цитате Дж. Солтон признает, может быть, сам того же желая, что в основе определения эффективности информационного поиска лежит понятие релевантности и что это понятие субъективно. К сказанному нужно лишь добавить, что данное понятие не перестает быть субъективным и тогда, когда суждение о релевантности документа выносится не одним потребителем, а судейской группой. Поэтому М. Таубе был совершенно прав, когда протестовал против трактовки более или менее субъективных оценок релевантности как некоей объективной характеристики.

Л. Дойл в принципе разделяет точку зрения М. Таубе и также не считает, что релевантность — адекватный критерий для оценки эффективности документальных ИПС. Но он справедливо пишет: «Релевантность — это умственный костыль (thought-crutch); с ней мы еще можем нестрого мыслить о проблеме информационного поиска, а без нее (или чего-то лучшего) мы совсем бы не могли

мыслить об этом. Понятие релевантности приносит больше пользы, чем вреда, если мы не принимаем его слишком буквально; оно позволяет нам в некоторой степени оптимизировать (to suboptimize) работу машинной части в системе «человек—машина» [88]. Таким образом, понятие релевантности пока все же остается единственной основой для любых оценок эффективности документальных ИПС.

За последние 10 лет для оценки эффективности документальных ИПС было предложено более десятка разных показателей. Девять из них оценивают лишь техническую эффективность ИПС,

| | $R$ | $\bar{R}$ | |
|---|---|---|---|
| $A$ | $a$ | $b$ | $a+b$ |
| $\bar{A}$ | $c$ | $d$ | $c+d$ |
| | $a+c$ | $b+d$ | $a+b+c+d$ |

Рис. 35. Таблица сопряженности релевантности и выдачи

$A$ — выданные документы,
$\bar{A}$ — невыданные документы,
$R$ — релевантные документы,
$\bar{R}$ — нерелевантные документы

т. е. такие характеристики, как точность и полнота информационного поиска. Все эти показатели тем или иным путем выведены из таблицы сопряженности релевантности и выдачи, приведенной на рис. 35. Пользуясь этой таблицей, можно также определить (разумеется, с определенной степенью приближения) различные показатели технической эффективности документальных ИПС в терминах теории вероятностей. Тогда:

$\frac{a}{a+c} = P_R(A)$ — условная вероятность выдачи релевантных документов

$\frac{b}{b+d} = P_{\bar{R}}(A)$ — условная вероятность выдачи нерелевантных документов

$\frac{c}{a+c} = P_R(\bar{A})$ — условная вероятность невыдачи релевантных документов

$\frac{d}{b+d} = P_{\bar{R}}(\bar{A})$ — условная вероятность невыдачи нерелевантных документов.

$$P_R(A) + P_R(\bar{A}) = 1$$

$$P_{\bar{R}}(A) + P_{\bar{R}}(\bar{A}) = 1$$

У. Буерн, Дж. Петерсон, Б. Лефковиц и Д. Форд (1961 г.) предложили оценивать общую эффективность документальной ИПС единой мерой, которая аналитически может быть определена так:

$$M = \sum_{i=1}^{i=n} w_i B_i \quad (n = 10,\ 11 \text{ или } 12),$$

[88] L. B. Doyle. Is relevance an adequate criterion in information system evoluation? В кн.: «Automation and scientific communication». Pt. 2. Washington, D. C., American Documentation Institute, 1963, стр. 200.

где $B_i$ — степень соответствия ИПС $i$-тому требованию, предъявленному к ней потребителем, $w_i$ — «весовой» коэффициент относительной важности $i$-того требования потребителя. В число предъявляемых к ИПС требований могут быть включены такие, как количество невыданных релевантных документов, величина «поискового шума», продолжительность поиска, сложность логики поиска, форма выдачи и т. д. Практического применения эта мера пока не получила [89].

Р. Уиллис (1962 г.) предложил, чтобы общая эффективность документальной ИПС, работающей в режиме поиска по методу последовательных приближений, на $k$-том шаге поиска оценивалась произведением четырех переменных:

а) отношения числа документов, выданных системой на $k$-том шаге поиска, к числу документов, выданных ею на шаге $(k-1)$;

б) стоимости использования ИПС (с учетом таких факторов, как затраты времени и средств на поиск, удобство поиска и стоимость индексирования);

в) числа невыданных релевантных документов;

г) штрафа за невыдачу релевантных документов (с учетом степени их релевантности) [90].

Мера для количественной оценки эффективности документальной ИПС, предложенная Дж. Верхуфом, У. Гоффманом и Дж. Белцером (1961 г.), может быть определена в символах таблицы сопряженности (см. рис. 35) следующим образом:

$$M = V_1 a - K_1 b - K_2 c + V_2 d,$$

где $V_1$ — ценность выдачи релевантных документов, $V_2$ — ценность невыдачи нерелевантных документов, $K_1$ — штраф за выдачу нерелевантных документов, $K_2$ — штраф за невыдачу релевантных документов. $V_1, V_2, K_1, K_2$ — неотрицательные постоянные величины. Не исключено, что некоторые из этих величин могут быть равны нулю.

Предполагается, что релевантность выданных документов определяют сами потребители. Однако разные потребители могут по-разному оценивать релевантность документов, выданных системой в ответ на один и тот же информационный запрос (рис. 36). Поэтому Дж. Верхуф и его соавторы с целью максимизации $M$ вводят для каждого документа, релевантного по отношению к данному запросу, понятие так называемой *критической вероятности* $P_c$, которая определяется так:

$$P_c = \frac{K_1 + V_2}{V_1 + K_1 + K_2 + V_2}.$$

[89] J. A. Swets. Information retrieval system. «Science», 1963, v. 141, № 3577, стр. 245.

[90] Там же, стр. 246.

Все документы, для которых вероятность их релевантности по отношению к информационному запросу больше $P_c$, подлежат выдаче в ответ на этот запрос [91].

К. Муэрс (1959 г.) предложил оценивать техническую эффективность документальной ИПС тремя следующими показателями:

а) отношением числа выданных высокорелевантных документов к общему числу таких документов в поисковом массиве;

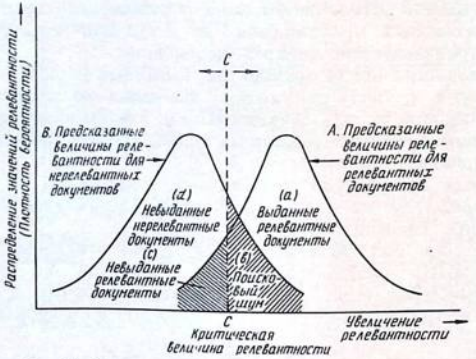



Рис. 36. Релевантность документов по отношению к информационному запросу

б) отношением числа выданных высокорелевантных и релевантных документов к общему числу таких документов в поисковом массиве;

в) отношением числа выданных нерелевантных документов к общему числу таких документов в поисковом массиве [92].

К. Муэрс считает, что для хорошей ИПС показатель (а) должен быть высоким, а показатель (в) — низким. Эффективность ИПС хороша, когда (а) близок к 1,0 и очень хороша, когда (в) меньше 0,01. Показатель (б) играет меньшую роль в определении эффективности ИПС; его величина всегда должна быть близка к 1,0.

---

[91] J. Verhoeff, W. Goffman and J. Belzer. Inefficiency of the use of Boolean functions for information systems. «Communications of the Association for Computing Machinery», 1961, v. 4, № 13, стр. 557—559.

[92] C. N. Mooers. The intensive sample test for the objective evaluation of the performance of information retrieval systems. RADC-TN-59-160 (ZTB-132). Cambridge, Mass., Zator Co., 1959.

Дж. Перри, А. Кент и М. Берри (1955 г.) [93] предложили в качестве основных показателей технической эффективности ИПС считать *коэффициент полноты* (recall factor) и *коэффициент точности* (pertinency factor) поиска, которые они определили следующим образом:

$$\text{Коэффициент полноты} = \frac{w}{x}, \text{ т. е. } \left[\frac{a}{a+c}\right].$$

$$\text{Коэффициент точности} = \frac{w}{m}, \text{ т. е. } \left[\frac{a}{a+b}\right],$$

где $w$ — число релевантных документов в выдаче $[a]$, $x$ — общее число релевантных документов в поисковом массиве $[a+c]$, $m$ — общее число релевантных и нерелевантных документов в выдаче $[a+b]$.

Такие же показатели были позже использованы С. Клевердоном (1961 г.) для эксперимента по оценке сравнительной эффективности четырех систем индексирования — алфавитно-предметной классификации, Универсальной десятичной классификации, фасетной классификации и системы унитермов. Это исследование получило название «Крэнфилдский проект» (из-за того, что оно проводится в местечке Крэнфилд, графство Бедфордшир, Великобритания). Исследование было начато в 1957 г. и продолжается до настоящего времени.

С. Клевердон [94] определяет *коэффициент полноты* (recall ratio) как выраженное в процентах отношение числа выданных релевантных документов к их общему числу в поисковом массиве $\left[\frac{a}{a+c}\right]$, а *коэффициент точности* (precision ratio, relevance ratio) — как выраженное в процентах отношение релевантных документов в выдаче к общему числу выданных релевантных и нерелевантных документов $\left[\frac{a}{a+b}\right]$. В аналитическом виде эти коэффициенты выражаются так:

$$R_r = \frac{100A}{C},$$

и

$$P_r = \frac{100A}{L},$$

где $R_r$ — коэффициент полноты, $P_r$ — коэффициент точности, $A$ — число релевантных документов в выдаче $[a]$, $C$ — общее число релевантных документов в поисковом массиве $[a+c]$,

[93] J. Perry, A. Kent and M. Berry. Machine literature searching. New York, Interscience Publishers, Inc., 1956, стр. 42—49.

[94] C. W. Cleverdon and J. Mills. The testing of index language devices. «Aslib Proceedings», 1963, v. 15, № 4, стр. 106—130.

$L$ — общее число релевантных и нерелевантных документов в выдаче $[a+b]$.

Общая зависимость полноты выдачи от точности информационного поиска показана на рис. 37. Более подробно коэффициенты точности и полноты С. Клевердона будут рассмотрены в конце данного раздела.

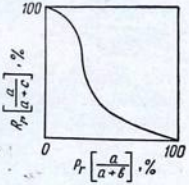



Рис. 37. Общая зависимость полноты выдачи $R_r\left[\frac{a}{a+c}\right]$ от точности информационного поиска $P_r\left[\frac{a}{a+b}\right]$ (по С. Клевердону)

Д. Суонсон (1960 г.) использовал для оценки технической эффективности ИПС показатель, который он определил следующим образом [95]:

$$M = R - pI,$$

где $R$ — сумма «весовых» индексов релевантности всех выданных документов, деленная на сумму «весовых» индексов релевантности (для данного информационного запроса) всех документов поискового массива; $I$ — эффективное количество нерелевантных документов ($I=N-RL$, где $N$ — общее число документов в выдаче $[a+b]$ и $L$ — общее число релевантных документов в поисковом массиве $[a+c]$); $p$ — коэффициент штрафа (penalty) за выдачу нерелевантных документов, принимающий разные численные значения в пределах от 0,05 до 0,15. Для идеальной ИПС, которая выдает все релевантные документы и при этом не выдает нерелевантных документов, $M=1$.

Г. Борко (1961 г.) [96] несколько модифицировал показатель Д. Суонсона, определив его так:

$$M = R - I,$$

где $I$ — отношение числа выданных нерелевантных документов к общему числу выданных релевантных и нерелевантных документов $\left[\frac{b}{a+b}\right]$.

В другой своей работе Д. Суонсон (1962 г.) предложил использовать для оценки технической эффективности ИПС такие пока-

[95] D. S. Swanson. Searching natural language text by computer. «Science», 1960, v. 132, № 3, стр. 1102—1103.

[96] J. A. Swets. Information retrieval systems. «Science», 1963, v. 141, № 3577, стр. 246.

затели, как коэффициент полноты поиска $\left[\frac{a}{a+c}\right]$ и «поисковый шум», измеряемый числом нерелевантных документов в выдаче $[b]$. Он построил график зависимости полноты выдачи $\left[\frac{a}{a+c}\right]$ от величины «поискового шума» $[b]$, который приведен на рис. 38.[97]

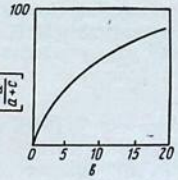



Рис. 38. Идеализированная кривая зависимости полноты выдачи $\left[\frac{a}{a+c}\right]$ от «поискового шума» $[b]$ (по Д. Суонсону)

Дж. Суетс (1963 г.) считает более правильным использовать для оценки величины «поискового шума» ИПС не абсолютное число нерелевантных документов в выдаче $[b]$ (Д. Суонсон) и не относительное количество релевантных документов в выдаче $\left[\frac{a}{a+b}\right]$ (Дж. Перри и др. и С. Клевердон), а отношение выданных нерелевантных документов к общему числу таких документов в поисковом массиве $\left[\frac{b}{b+d}\right]$. При таком подходе охватывается вся информация, содержащаяся в таблице сопряженности (см. рис. 35). Кроме

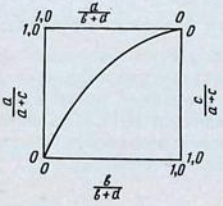



Рис. 39. Типовая рабочая характеристическая кривая ИПС (по Дж. Суетсу)

того, для этих показателей теория статистических решений позволяет построить семейство кривых, аналогичных эмпирической кривой Д. Суонсона. Характеристическая кривая такого типа приведена на рис. 39 [98].

У. Гоффман и В. Ньюилл (1964 г.) решили использовать для оценки технической эффективности документальной ИПС показа-

[97] D. S. Swanson. Interrogating a computer in natural language. В кн.: «Information processing 1962; Proceedings of the IFIP Congress 62». Amsterdam, North Holland Publishing Co., 1963, стр. 288—292.

[98] J. A. Swets. Information retrieval systems. «Science», 1963, v. 141, № 3577, стр. 246—250.

исследования, проведенного в 1963 г. специалистами американской фирмы «Артур Д. Литтл, инк.» на массиве 70 тыс. документов [101].

Таким образом, когда в какой-то ИПС достигнуто оптимальное соотношение $R_r$ и $P_r$, то дальнейшее увеличение $R_r$ возможно только за счет уменьшения $P_r$, и наоборот. Общая зависимость $R_r$ от $P_r$ показана на рис. 37. Участники Крэнфилдского проекта С. Клевердон, Ф. Ланкастер и Дж. Миллз [102] считают, что в пределах рабочих характеристик ИПС по физике, технике и металлургии увеличение точности информационного поиска на 1% ведет к уменьшению его полноты на 3%. Они же утверждают [103], что в современных ИПС коэффициенты полноты поиска обычно составляют от 70 до 90%, а коэффициенты точности — от 8 до 20%.

Выбор соотношения показателей полноты и точности поиска должен определяться теми задачами, для решения которых создается ИПС. Для удовлетворения информационных потребностей исследователей и конструкторов едва ли необходима полнота поиска, превышающая 85—90%. В данном случае важней сохранить сравнительно высокую точность поиска. Когда же полнота поиска достигает 90%, то ее дальнейшее увеличение, как это видно из рис. 37, сопровождается значительно более быстрым падением точности поиска. Не будет преувеличением сказать, что за обеспечение выдачи последних 10% релевантных документов приходится платить такую же, а может быть, и бóльшую цену, как и за обеспечение выдачи предшествующих 90% таких документов. С другой стороны, если ИПС предназначается, например, для использования при экспертизе патентных заявок на новизну, то решающее значение приобретает именно полнота поиска.

Ранее уже отмечалось, что применяемые С. Клевердоном критерии не учитывают такого важного параметра ИПС, как объем поискового массива. Между тем приемлемость того или иного коэффициента точности зависит от среднего объема выдачи, с увеличением которой повышается минимально приемлемый коэффициент точности. Но средний объем выдачи находится в прямой зависимости от общего объема поискового массива. Таким образом, рост поискового массива требует увеличения точности информационного поиска. При этом необходимо также обеспечить достаточно большую полноту поиска. Справедливость данного вывода можно доказать путем следующего несложного расчета.

---

[101] Centralization and documentation. Final report to the National Science Foundation. C—64469. Cambridge, Mass., Little, 1963, стр. 40.

[102] C. W. Cleverdon, F. W. Lancaster and J. Mills. Uncovering some facts of life in information retrieval. «Special Libraries», 1964, v. 55, № 2, стр. 88.

[103] C. Cleverdon. The Cranfield hypothesis. «Library Quarterly», 1965, v. 35, № 2, стр. 121—122.

Допустим, что техническая эффективность некоторой документальной ИПС характеризуется следующими показателями: $R_r=70\%$ и $P_r=16\%$. Предположим, также, что средний объем выдачи ($L$) составляет 0,5% общего поискового массива (обозначим его через $S$). Вычислим абсолютные значения числа релевантных документов в выдаче ($A$), числа невыданных релевантных документов ($C-A$) и $L$ (в целых числах) при разных значениях $S$ (табл. 20).

Таблица 20

**Зависимость характеристик выдачи от общего объема поискового массива**

| $S$ | $A$ | $C-A$ | $L$ |
|---|---|---|---|
| $10^4$ | 8 | 3 | 50 |
| $10^5$ | 80 | 34 | 500 |
| $10^6$ | 800 | 343 | 5000 |

Из этих цифр видно, что уже при $S=10^4$ средний объем выдачи составляет 50 документов, из которых только 8 релевантны. Следовательно, потребитель должен путем последовательного перебора этих 50 документов найти среди них восемь документов, которые предположительно отвечают на его информационный запрос. И если это может быть приемлемым для потребителя при $S=50$, то, безусловно, неприемлемо при $S=500$. Необходимость отыскивать 80 релевантных документов среди 500 выданных означает, что ИПС в значительной мере уже перестает выполнять свою функцию. Сделанный вывод в принципе не зависит от среднего объема $L$. Если бы мы приняли $L=0{,}1\%$, то описанная выше ситуация возникла бы не при $S=10^5$, а при $S=5\cdot10^5$.

Аналогичная картина наблюдается и в отношении полноты выдачи. Невыдача 3 релевантных документов из 11 при $S=10^5$ может быть терпимой, но едва ли терпима невыдача 34 документов из 114 ($S=10^5$).

Глава 5

# ТРАДИЦИОННЫЕ ИНФОРМАЦИОННО-ПОИСКОВЫЕ ЯЗЫКИ

Прежде, чем перейти к общей оценке эффективности библиотечно-библиографических и алфавитно-предметных классификаций как семантических систем, предназначенных для поиска документов, примем в качестве исходных два следующих тезиса:

1. Информационный поиск в систематическом или алфавитно-предметном каталоге (указателе) производится не их создателями, а самими потребителями.

2. Отыскание какого-либо предмета в множестве других предметов возможно лишь тогда, когда выполнено хотя бы одно из следующих двух условий:

а) известны система пространственного расположения предметов относительно друг друга и место (координаты) искомого предмета в этой системе;

б) известны отличительные признаки искомого предмета, позволяющие идентифицировать этот предмет во множестве других предметов путем последовательного перебора всех элементов этого множества (или, по крайней мере, определенного подмножества таких элементов).

Очевидно, что при достаточно большом числе элементов в этом множестве (порядка $10^3$ и более) их последовательный перебор вручную практически невозможен. Но систематические и алфавитно-предметные каталоги и указатели рассчитаны именно на индивидуальное использование без применения каких-либо средств механизации. Поэтому для таких каталогов (указателей) особое значение приобретает условие 2а, т. е. требование упорядоченности элементов множества, в котором производится информационный поиск.

Для рассмотрения структуры и основных свойств библиотечно-библиографических и алфавитно-предметных классификаций введем следующие определения:

*Предмет* — это все, что может быть объектом специального изучения. Предметами являются не только материальные вещи

объективного мира с их свойствами и отношениями, но и отражения этих вещей в нашем сознании, т. е. ощущения, восприятия, представления, понятия, суждения, умозаключения, научные теории и т. д. В логике такие объекты называются *предметами мысли*[1]. Предмет может быть отраслевым, когда он является объектом изучения или применения лишь в одной отрасли науки или техники, и межотраслевым, когда он изучается или применяется в нескольких отраслях науки и/или техники.

*Предмет документа* (*информационного запроса*) — это основная тема (одна или несколько), которой посвящен данный документ (информационный запрос).

*Предметизация* — это процедура определения и формулирования предметов документа или информационного запроса.

*Предметизатор* — это лицо, осуществляющее предметизацию.

*Систематизация документов* — это распределение документов по классам некоторой классификационной системы на основании смыслового содержания или других отличительных признаков этих документов. Синоним этого термина — *классифицирование документов*.

*Систематизатор* — лицо, осуществляющее систематизацию (классифицирование) документов.

*Тематический запрос* — это информационный запрос, который охватывает не один, а некоторое множество (класс) предметов, связанных с предметом запроса отношениями «вид—род», «часть—целое» и т. д.

*Узкопредметный запрос* — это информационный запрос по единичному предмету, т. е. такому, который не имеет видовых подразделений.

Некоторые исследователи считают, что предметами документа могут быть любые сведения, которые в явном или скрытом виде содержатся в документе. С такой точкой зрения нельзя согласиться. Как правило, документ—статья, научно-технический отчет, патентное описание и т. д. — посвящен рассмотрению всего одной или нескольких тем. Автор документа обычно стремится раскрыть эту тему (или темы) в заглавии, которое он дает своему документу. Задача предметизатора состоит в том, чтобы определить и должным образом сформулировать эту основную тему (или темы).

Если документ (например, монография по основам информационного поиска) охватывает широкий круг вопросов, то такой документ можно рассматривать как совокупность отрезков текста (микродокументов), каждый из которых посвящен раскрытию лишь одной или нескольких основных тем.

С другой стороны, любой документ может быть объектом соответствующего исследования, в результате которого специалист иногда извлекает из текста документа даже такие сведения,

[1] Д. П. Горский. Логика. Изд. 2. М., Учпедгиз, 1963, стр. 30.

которые в явном виде в нем не содержатся. Очевидно, что побочные сведения нельзя считать предметом такого документа.

Каждый документ охватывает неопределенно большое число предметов, причем некоторые из этих предметов могут в нем лишь упоминаться или подразумеваться. Поэтому если не проводить различия между главными и второстепенными предметами документа, то тогда понятие «предмет документа» потеряет свою смысловую инвариантность и станет чисто субъективным представлением, которое будет определяться специальностью, уровнем квалификации и научными интересами предметизатора.

## ОБЩИЕ ПРИНЦИПЫ КЛАССИФИКАЦИИ

Классификацией называется распределение предметов или отношений по классам на основании общего признака, присущего предметам данного рода и отличающего их от предметов других родов, причем распределение производится с таким расчетом, чтобы каждый класс относительно других классов занимал в получившейся системе определенное, точно фиксированное место. В основе классификации лежит деление понятия. Признак, по которому производится деление, называется *основанием деления* и является переменным [2].

Классификация должна подчиняться следующим формально-логическим правилам:

1) классификация должна производиться только по одному основанию;

2) получаемые в результате деления подклассы должны исключать друг друга;

3) деление на подклассы должно быть соразмерным;

4) деление на подклассы должно быть непрерывным, без скачков.

Пусть $K$ — некоторый класс, который на основе соответствующего признака разделен на подклассы $k_1, k_2, k_3, \ldots, k_n$. Тогда сформулированные выше правила в обобщенном виде можно записать следующим образом:

1) $k_1 \cup k_2 \cup k_3 \cup \ldots \cup k_n = K$

2) $k_1 \cap k_2 \cap k_3 \cap \ldots \cap k_n = \varnothing$

3) $k_1 \cap k_2 = \varnothing$, $k_2 \cap k_3 = \varnothing$, $k_1 \cap k_3 = \varnothing$ и т. д. — для любой пары подклассов.

Основные виды отношений между подразделениями классификации — это *иерархия* (*подчинение*) и *соподчинение*. Иерархия — это такое, и только такое отношение, при котором один класс является подклассом другого, более широкого. Отношением со-

[2] Г. К л а у с. Введение в формальную логику. М., ИЛ, 1960, стр. 254.

подчинения связаны классы, которые суть подклассы одного, более широкого класса. Иерархия имеет место между классом $K$ и подклассом $k_i$ ($k_i \subset K$, $i=1, 2, 3, \ldots, n$), а соподчинение — между подклассами $k_1$, $k_2$, $k_3$, . . ., $k_n$ ($k_1 \subset K$, $k_2 \subset K$, $k_3 \subset K$ и т. д.). Иногда различают два вида иерархии: сильную и слабую. При сильной иерархии каждый подкласс имеет один, и только один, непосредственно предшествующий ему класс, при слабой — более одного класса. Классификации, между подразделениями которых существуют лишь отношения сильной иерархии и соподчинения, будем далее называть *иерархическими классификациями*. Следует сразу же оговориться, что иерархические классификации в чистом виде существуют только как формальные структуры, оторванные от действительности.

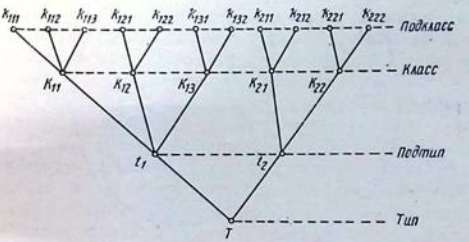



Рис. 40. «Дерево» иерархической классификации

Наиболее удобной формой изображения иерархических классификаций считаются графы типа «деревьев». *Граф* — это фигура, состоящая из точек (называемых *вершинами*) и отрезков, соединяющих некоторые из этих вершин. Эти соединительные отрезки, называемые *ребрами*, могут быть прямолинейными и криволинейными. Граф, каждая вершина которого соединена определенной цепью с любой другой его вершиной, но который при этом не содержит замкнутых цепей (*циклов*), называется *деревом*. Любая вершина, выбранная за начальную точку «дерева», называется его *корнем* [3]. Пример такого «дерева» показан на рис. 40.

Для наглядного изображения отношений между объемами понятий (классами) применяются круги Эйлера (в отличие от диаграмм Венна, применяемых для изображения о п е р а ц и й над классами). На рис. 41 при помощи кругов Эйлера показаны отношения между подразделениями в иерархической классификации, «дерево» которой было дано на рис. 40.

[3] О. О р е. Графы и их применение. М., «Мир», 1965, стр. 47—52, 168—172.

Классификации могут быть *естественные* и *искусственные*. В естественных классификациях в качестве основания деления используются существенные признаки, свойства самих предметов, тогда как в искусственных классификациях исходят обычно из несущественных признаков. Но резкой границы между искусственными и естественными классификациями провести нельзя. Вот что пишет по этому поводу известный логик Г. Клаус (ГДР):

«Признаки, которые, вообще говоря, несущественны, в определенных условиях могут оказаться существенными. Так, например, геометрические размеры и вес машин, приборов и т. д.

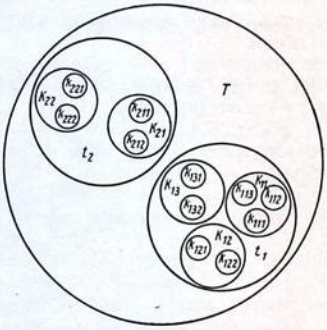



Рис. 41. Отношения между подразделениями иерархической классификации, изображенной на рис. 40

в общем не являются существенными признаками. При установлении же платы за провоз по железной дороге важно именно это. Вот почему указанные признаки, сами по себе несущественные, используются здесь в качестве основания классификации для составления тарифной сетки» [4].

Искусственные классификации не имеют познавательного значения. Они создаются лишь с целью систематизации предметов. Поэтому в качестве оснований деления в них выбираются признаки, удобные для этой цели, но не относящиеся к категории существенных.

Д. П. Горский предлагает делить все классификации на *естественные* и *вспомогательные*, причем вспомогательными классификациями он считает такие, которые создаются «с целью наиболее легкого отыскания того или иного индивидуального предмета

[4] Г. Клаус. Введение в формальную логику. М., ИЛ, 1960, стр. 254.

среди других классифицируемых предметов» [5], и призывает отличать их от классификаций искусственных. В качестве примера вспомогательной классификации Д. П. Горский называет алфавитное расположение слов в словаре. При вспомогательной классификации знание места классифицируемого предмета в системе не дает возможности что-либо утверждать о свойствах этого предмета. Этим вспомогательная классификация отличается от классификации искусственной. Но такая точка зрения не является общепризнанной.

Большинство классификаций — как естественных, так и искусственных — всегда приблизительно, на что неоднократно указывал Ф. Энгельс.

«*Hard and fast lines* [*абсолютно резкие разграничительные линии*] несовместимы с теорией развития. Даже разграничительная линия между позвоночными и беспозвоночными уже более не безусловна, точно также между рыбами и амфибиями; а граница между птицами и пресмыкающимися с каждым днем все более и более исчезает. . . Диалектика, которая точно так же не знает hard and fast lines [абсолютно резких разграничительных линий] и безусловного, пригодного повсюду «или — или», которая переводит друг в друга неподвижные метафизические различия, признает в надлежащих случаях наряду с «или — или» также «как то, так и другое» и опосредствует противоположности — является единственным, в высшей инстанции, методом мышления, соответствующим теперешней стадии развития естествознания. Разумеется, для повседневного обихода, для научной мелкой торговли метафизические категории сохраняют свое значение» [6].

Характеризуя биологические исследования, систематически проводившиеся с середины XVIII в., Ф. Энгельс писал:

«Чем глубже и точнее велось это исследование, тем больше расплывалась под руками. . . застывшая система неизменно установившейся органической природы. Не только все более и более расплывчатыми становились границы между отдельными видами растений и животных, но и обнаружились животные, как ланцетник и чешуйчатник, которые точно издевались над всей существовавшей до того классификацией» [7]. «Понятия . . . вид, род, класс, благодаря теории развития стали текучими и тем самым *относительными*» [8].

Классификации использовались для систематизации книг и других документов по их содержанию уже в глубокой древности. Известно, что в одной из древнейших библиотек — в Ниневийской

---

[5] Д. П. Горский. Логика. Изд. 2. М., Учпедгиз, 1963, стр. 86.
[6] Ф. Энгельс. Диалектика природы. М., Госполитиздат, 1955, стр. 167.
[7] Там же, стр. 10—11.
[8] Там же, стр. 180.

библиотеке ассирийского царя Ашурбанипала (VII в. до н. э.) — клинописные глиняные плитки систематизировались в соответствии с их содержанием («история», «закон», «переписка» и т. д.) [9].

Классификации, специально предназначенные для систематизации книг и других документов, называются *библиотечно-библиографическими классификациями* (ББК). «Библиотечно-библиографическая классификация, — пишет Е. И. Шамурин, — классификация, приспособленная для библиотечно-библиографических нужд, т. е. для размещения произведений печати на полках, расстановки их описаний в систематическом каталоге и расположения материалов в библиографических указателях» [10]. Эти искусственные классификации обладают рядом специфических свойств, которые отличают их как от формально-логических классификаций, так и от классификаций наук.

Основной целью естественных классификаций является установление и обозначение существенных связей между предметами. Задача ББК совершенно иная: ББК должна обеспечить такую систематизацию множества документов, при которой документы по одному и тому же предмету (или описания этих документов) будут физически собираться в одном, и только в одном, месте. Таким образом, для ББК исключительно большое значение имеет ясное обозначение относительного расположения классов в схеме, так как только благодаря этому можно:

а) помещать предмет (книгу, статью, реферат и т. д.), не имеющий достаточно краткого индивидуального имени (под именем в данном случае понимается текст, выражающий основную тему какого-либо документа), в соответствующее место схемы или раскрывать содержание этого предмета через его положение в классификационной схеме;

б) быстро находить предмет (книгу, статью, реферат и т. д.), если известно, какое место закреплено за ним в схеме;

в) производить группировку таких предметов, которые с точки зрения потребителя предпочтительно иметь в одном месте.

«Классификация книг не исчерпывается распределением их по отделам, — писал З. Н. Амбарцумян. — Различие книг по их содержанию обусловливает образование большого количества отделов. Чтобы легко и точно ориентироваться в этих отделах и быстро находить нужный отдел или его подразделения, необходимо установить определенную четкую их последовательность. Это достигается систематическим порядком расположения отделов» [11].

---

[9] З. Н. Амбарцумян. Библиотечная классификация. М., Госкультпросветиздат, 1947, стр. 28.

[10] Е. И. Шамурин. Словарь книговедческих терминов. М., «Сов. Россия», 1958, стр. 33.

[11] З. Н. Амбарцумян. Указ. соч., стр. 7.

В формальной логике классификация представляет собой метод, в известной степени заменяющий определение[12], когда последнее по тем или иным причинам невозможно. Это означает, что для формальной логики важно прежде всего образование классов предметов и установление *родо-видовых отношений* (иерархии) между этими классами, а не их пространственное расположение относительно друг друга. Логическое отношение включения лишь определяет, что подклассы необходимо расположить таким образом, чтобы было показано их подчинение классу. Но формальная логика не дает нам универсального метода построения классов — даже в сравнительно узкой области, — которые бы были связаны только родо-видовыми отношениями. Кроме того, она не может подсказать нам, в каком порядке следует располагать элементы деления. Например, если мы разделим класс «математика» на подклассы «арифметика», «алгебра», «геометрия», «тригонометрия» и т. д., то порядок этих подклассов в ряду определяется содержательно, а не по формально-логическим правилам. Таким образом, ББК отличается от формально-логической классификации своим принципом упорядоченного размещения классов в пространстве. Это различие имеет принципиальное значение, и оно не должно упускаться из виду.

Итак, формальная логика не дает ответа, в каком порядке следует располагать классы в системе. Поэтому приходится обращаться к классификациям наук. Выяснение отношения ББК к классификации наук вообще имеет большое принципиальное значение. Если рассматривать этот вопрос в историческом плане, то тесная связь ББК с классификациями наук очевидна. Более того, некоторые из первых классификаций наук начали создаваться в значительной мере под воздействием практической потребности библиотек в систематизации книг[13].

Во второй половине XIX в. работы по созданию различных схем ББК значительно активизировались. Это можно объяснить быстрым ростом библиотечных фондов в связи с огромными успехами науки и техники, а также увеличением значения справочно-библиографической работы. Именно в этот период и были созданы «Десятичная классификация» М. Дьюи (1876 г.) и «Растяжимая классификация» Ч. Э. Кеттера (1882 г.), которые оказали влияние на все последующие ББК. Создатели этих двух классификаций занимали противоположные позиции при определении связи ББК

---

[12] «Определение есть логический прием, позволяющий: а) отличать, отыскивать, строить интересующий нас предмет; б) уточнять значение уже введенного в науку термина, а также формировать значение вновь вводимого термина». См.: Д. П. Горский. Логика. Изд. 2. М., Учпедгиз, 1963, стр. 67.

[13] См.: Б. М. Кедров. Классификация наук, т. 1. М., Изд-во ВПШ и АОН при ЦК КПСС, 1961, стр. 43—44.

с классификациями наук. Е. И. Шамурин писал по этому поводу: «В противоположность Дьюи, не придававшему значения научности библиотечной классификации, считавшему, что читатель может легко найти нужный ему материал по алфавитно-предметному указателю, Кеттер был убежден, что лучшей практической классификацией является та, которая разработана на научной основе» [14].

Такое положение в значительной мере сохранилось и до настоящего времени. Оно объясняется противоречивостью требований, предъявляемых к ББК. С одной стороны, ББК должна соответствовать классификации наук, так как она предназначается для систематизации письменных документов, в которых излагаются результаты научного познания объективного мира.

«Книги — передатчики человеческого знания и опыта, накопленного за тысячелетия, в книгах выражается познание мира, окружающей нас действительности», — пишет А. В. Кленов. «Если возможно, как это делает философия, разрозненные знания сводить в единую стройную систему классификации наук или, как говорят, «карту знаний», то тем самым возможно строить и соответствующую ей систему классификации книг, распределяя книги по содержанию.

Однако полного совпадения здесь быть не может» [15].

С другой стороны, ББК должна строиться по формально-логическим правилам, так как только при этом условии она может удовлетворительно выполнять функцию информационно-поискового языка, т. е. однозначно определять (специфицировать) публикации. Но это второе требование не может быть предъявлено к классификации наук, так как тогда классификация получается «искусственной, формальной, основанной на допущении резких границ между науками, которых в действительности не существует» [16]. Принцип развития, положенный Ф. Энгельсом в основу классификации наук, исключает признание каких-либо резких, абсолютных разрывов между науками. «При указанном подходе, — пишет Б. М. Кедров, — становились уже неприемлемыми обычные формально-логические правила деления. В общем классификационном ряду оказывались такие участки, где члены ряда не исключали друг друга, а проникали один в другой, образуя переходы, «мосты» между собой. Такие участки нельзя было целиком включить в одну из двух смежных, соприкасающихся между собой наук, например, *либо* только в физику, *либо* только в химию,

---

[14] Е. И. Шамурин. Очерки по истории библиотечно-библиографической классификации, т. 2. М., Всесоюз. кн. палата, 1959, стр. 192.

[15] А. В. Кленов. Библиотечные каталоги. М., Всесоюз. кн. палата. 1963, стр. 174.

[16] Б. М. Кедров. Классификация наук, т. 1. М., Изд-во ВПШ и АОН при ЦК КПСС, 1961, стр. 23.

а их следовало относить одновременно *и* к той *и* к другой науке. Здесь необходимо применение диалектической логики» [17].

Эту же важную мысль Б. М. Кедров подчеркивал в другой своей работе. «Для *формальных* классификаций, — писал он, — характерны принцип координации и принцип искусственности, составляющие вместе логическую основу этих классификаций» [18]. И далее: «Логическая сущность принципа координации состоит в последовательном проведении положений формальной логики, в частности ее требования, чтобы члены деления взаимно и полностью исключали друг друга. Это возможно при условии, когда науки резко обособляются одна от другой и соотносятся лишь внешним образом» [19].

Из сказанного выше следует, что между классификацией наук и ББК имеется существенное различие. Но это различие не такое, которое бы заставляло строить ББК в полном отрыве от классификации наук. Насколько это возможно, ББК должна строиться на базе классификации наук. Но это означает, что в такой ББК не могут быть полностью выполнены все требования, предъявленные к классификации формальной логикой.

Большинство существующих ББК следует «естественному порядку». Именно к этому необходимо стремиться, особенно при разработке ББК для естественных наук. Но имеются и такие классификации, в которых порядок расположения классов в ряду определяется чисто эмпирическими соображениями. Такова, например, классификация Библиотеки конгресса США. В фасетных классификациях критерием для определения предпочтительного порядка фасетов является практическое удобство.

Классификация, в том виде, в каком она используется в формальной логике, близка к процедуре описания предмета, его спецификации. Под спецификацией в данном случае понимается точное, однозначное и непротиворечивое обозначение предмета без указания его относительного положения в классификационной схеме. С. Р. Ранганатан определяет библиотечно-библиографическую классификацию как «. . . перевод названия предмета книги на более удобный (preferred) искусственный язык порядковых чисел и индивидуализацию нескольких книг, посвященных рассмотрению одного и того же предмета, при помощи следующего множества порядковых чисел, которые обозначают соответствующие признаки книги, не имеющие отношения к предполагаемому ее содержанию. Первое из этих порядковых чисел называется номером класса книги. Второе порядковое число — номером книги

---

[17] Б. М. Кедров. Классификация наук, т. 1. М., Изд-во ВПШ и АОН при ЦК КПСС, 1961, стр. 25.

[18] Б. М. Кедров. Формальные и диалектические принципы классификации наук и общая структура научного знания. В сб.: «Формы мышления». М., Изд-во АН СССР, 1962, стр. 273.

[19] Там же, стр. 274.

этого класса». И далее: «Библиотечная классификация — это обозначение (reduction) предметов и книг порядковыми числами, необходимое для достижения специфических целей» [20].

## ИЕРАРХИЧЕСКИЕ БИБЛИОТЕЧНО-БИБЛИОГРАФИЧЕСКИЕ КЛАССИФИКАЦИИ

Иерархические классификации, как уже было отмечено выше, с древнейших времен широко применяются для систематизации и поиска книг в библиотеках. В настоящее время наиболее распространенными из таких классификаций являются Десятичная классификация Дьюи, Библиографическая классификация Блисса и классификация Библиотеки конгресса США. Схемы этих классификаций и методика их использования широко описаны в литературе [21]. Поэтому здесь эти вопросы обсуждаться не будут. В данном разделе делается попытка показать наиболее существенные недостатки любой иерархической классификации как системы для поиска документов.

К иерархической ББК, предназначенной для целей информационного поиска, предъявляются следующие основные требования:

— для любого предмета или предметного комплекса в классификации должен быть предусмотрен один, и только один, исчерпывающий класс;

— классификация должна обеспечивать информационный поиск по любому сочетанию признаков.

Для удовлетворения этих требований в иерархической классификации необходимо иметь отдельные исчерпывающие классы для всех возможных предметов и предметных комплексов, т. е. перечислить их в классификационных таблицах. Поэтому такие классификации называются *перечислительными*. Однако ни в одной классификации невозможно заранее перечислить все предметы и предметные комплексы, так как их общее количество выражается астрономическими числами и с развитием науки и техники непрерывно возрастает. Кроме того, общие принципы построения иерархических классификаций таковы, что они не позволяют нам иметь в классификации один, и только один, исчерпывающий класс для каждого предмета или предметного комплекса. Справедливость этих утверждений можно доказать как путем формально-логического, так и путем содержательного анализа иерархических классификаций.

---

[20] S. R. Ranganathan. Elements of library classification. Bombey a. o, Asia Publ. House, 1962, стр. 10, 14.

[21] J. Mills. A modern outline of library classification. L., Chapman and Hall Ltd., 1960.

Начнем с формально-логического анализа иерархических классификаций. С этой целью построим графическую модель иерархической классификации.

Классификацию будем строить дедуктивно, т. е. путем последовательного уменьшения объема исходного понятия (summum genus). В качестве оснований деления используем признаки (характеристики) α, β и γ. При этом условимся, что в результате применения каждого признака к делимому классу мы получаем три подкласса. Условимся также, что прописными буквами $A$,

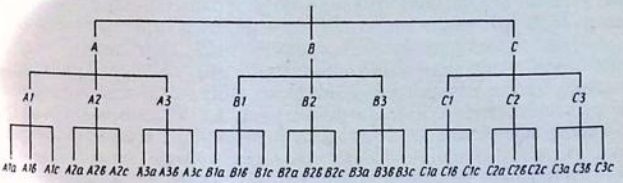



Рис. 42. Граф иерархической классификации

$B$ и $C$ будем обозначать классы первого уровня иерархии, получаемые в результате применения к исходному классу признака α; цифрами 1, 2 и 3 — подклассы, порождаемые признаком β в результате его применения к каждому классу первого уровня иерархии; строчными буквами $a$, $b$ и $c$ — подклассы, порождаемые признаком γ в результате его применения к каждому классу второго уровня иерархии. Тогда граф классификации будет иметь вид, показанный на рис. 42.

Допустим, что в качестве исходного класса в данном случае взят класс «гражданские самолеты», а в качестве оснований деления — соответственно следующие признаки: «целевое назначение» (α), «число моторов» (β) и «расположение крыла» (γ). Применение признака α к исходному классу дает нам подклассы «пассажирские самолеты» ($A$), «транспортные самолеты» ($B$) и «специальные самолеты» ($C$). Применение признака β соответственно к классам $A$, $B$ и $C$ дает нам подклассы «одномоторные пассажирские самолеты» ($A1$), «двухмоторные пассажирские самолеты» ($A2$), «четырехмоторные пассажирские самолеты» ($A3$) и т. д. Наконец, применение признака γ соответственно к классам $A1$, $A2$, $A3$, $B1$, $B2$, $B3$, $C1$, $C2$ и $C3$ порождает подклассы «одномоторные пассажирские самолеты с низким расположением крыла» ($A1a$), «одномоторные пассажирские самолеты со средним расположением крыла» ($A1b$), «одномоторные пассажирские самолеты с высоким расположением крыла» ($A1c$) и т. д.

При построении этой классификации деление, как того требуют формально-логические правила, должно производиться только по одному основанию. Если бы на первом шаге деления были применены одновременно два признака — например, «целевое назначение» (α) и «число моторов» (β), то мы получили бы на одном уровне иерархии пересекающиеся классы «пассажирские самолеты» (*A*) и «одномоторные самолеты» (1) и т. д. Тогда нам пришлось бы помещать документы об одномоторных пассажирских самолетах не в один, а в два класса.

Описанная классификация, безусловно, не является идеальной. При ее построении, по-видимому, следовало бы после признака «целевое назначение» (α) применить признак δ — «тип движителя» (воздушный винт, реактивный двигатель), а лишь после него — признак «число моторов» (β). Классификация не является также исчерпывающей. Например, в число подклассов, образованных путем применения признака «число моторов» (β), не входят подклассы «трехмоторные самолеты» и «шестимоторные самолеты». Из-за несоблюдения этих формально-логических правил построения классификации мы не находим в ней класса, к которому можно было бы отнести, например, документ о 3-моторных транспортных самолетах или о реактивных пассажирских самолетах. Следовательно, при построении иерархической классификации соблюдение установленных формально-логических правил имеет исключительно большое значение.

Однако соблюдение установленных формально-логических правил при построении иерархической классификации не устраняет главных недостатков такой классификации. Эти недостатки заключаются в том, что при использовании иерархической классификации мы не можем производить многоаспектное индексирование документов и информационный поиск по любому сочетанию характеристик.

Одно из основных требований, предъявляемых к ББК, состоит в том, чтобы она обеспечивала сбор документов по определенному предмету в одном, и только в одном, классе. Но это требование не выполняется даже в такой иерархической классификации, которая безупречна в формально-логическом отношении. Для построения иерархической классификации используется определенный ряд признаков (оснований деления). Число этих признаков обычно бывает невелико, и они применяются только в одной последовательности. Такая процедура позволяет построить классы предметов, обладающих лишь определенными сочетаниями признаков, выбранных в качестве оснований деления.

В классификации, которая приведена на рис. 42, содержатся классы, обладающие следующими сочетаниями признаков: α, αβ, αβγ, αγ, β, βγ, γ. Но в этой иерархической классификации исчерпывающие классы имеются только для трех из перечисленных сочетаний — α, αβ и αβγ. Если же эту классификацию ис-

пользовать для информационного поиска по остальным сочетаниям признаков, то для этого каждый раз придется собирать документы из нескольких классов. Например, документы о гражданских самолетах с высоким расположением крыла в нашей классификации распределены по 9 классам, документы о двухмоторных гражданских самолетах — по 3 классам и т. д. Можно легко подсчитать, что в иерархической классификации число классов, обладающих каким-то видовым признаком, равно произведению числа видовых признаков, которые используются для образования каждого из уровней иерархии, расположенных выше того уровня, где впервые появился интересующий нас признак.

Очевидно, что если нам потребуются документы, предмет которых выражен, например, сочетанием признаков $\alpha$, $\beta$ и $\delta$, где $\delta$ — признак, не использовавшийся при построении иерархической классификации в качестве основания деления, то эта классификация вообще не позволит найти нужные нам документы, так как в ней отсутствует класс, обладающий таким сочетанием признаков. При построении классификации, приведенной на рис. 42, в качестве основания деления не использовался признак «тип движителя». Поэтому эта классификация оказывается практически бесполезной при поиске документов, например, о транспортных реактивных самолетах.

Для обеспечения возможности быстрого поиска документов по любому сочетанию признаков, взятых из некоторого данного множества, а также для предотвращения рассеяния таких документов по разным классам, необходимо построить отдельные классы для всех возможных сочетаний этих признаков. Эффективность этого метода зависит от числа классов самого низшего уровня иерархии (обозначим это число через $N$), которое можно построить путем применения в разной последовательности данного множества признаков. Очевидно, что чем больше $N$, тем выше трудоемкость процедуры классифицирования документов.

Ориентировочная оценка трудоемкости этой процедуры может быть произведена путем определения порядка величины при минимальном числе признаков, используемых в качестве оснований деления, и минимальном числе подклассов, порождаемых каждым из этих признаков. В сущности, задача сводится к определению числа иерархий, которое можно построить путем применения данного множества признаков.

Обозначим через $C$ множество признаков, используемых в качестве оснований деления. $C=\{c_1, c_2, c_3, \ldots, c_n\}$, где $n$ — число элементов этого множества. Пусть $e$ — число классов, порождаемых каждым признаком, взятым из $C$. В реальных условиях $e$ — величина переменная; $e=f(c_i)$, где $c_i \in C$. Но для упрощения задачи будем считать, что $e$ — величина постоянная. Теперь предположим, что $n=3$ и $e=3$. Такая ситуация соответствует одному из самых простейших случаев построения иерархических класси-

фикаций и воспроизведена на рис. 42. В реальных ББК численные значения $n$ и $e$ будут значительно выше. Задача состоит в том, чтобы определить $N$ — число иерархий, которое можно построить при $n=e=3$.

Если обратиться к рис. 42, то из него видно, что общее число классов $N_i$, которое порождает однократное применение трех признаков при $e=3$, составляет $3^3$. Но при построении классификации мы можем выбрать ту или иную последовательность расположения классов на первом уровне иерархии. Число возможных перестановок классов $A$, $B$ и $C$ равно 3! Очевидно, что если мы выбрали определенную последовательность расположения классов на первом уровне иерархии, то этой последовательности мы будем придерживаться и на ее последующих уровнях. Это означает, что на последующих уровнях иерархии каждый класс непосредственно вышестоящего уровня порождает 3! подклассов. Таким образом, общее число классов $N_j$ на третьем, самом низшем уровне будет составлять $(3!)^3$. Однако путем многократного применения в разной последовательности признаков, взятых из $C$, можно получить 3! иерархий. Отсюда следует, что $N=(3!)^3 \times 3!=(3!)^4$. В общем виде $N=(n!)^{n+1}$.

Теперь можно определить численное значение $N$ при $n=3$. $N=(3!)^4=1296$. Полученный результат свидетельствует о том, что даже при явно заниженных значениях $n$ и $e$ численная величина $N$ настолько велика, что это делает предложенный прием повышения эффективности иерархической классификации для информационного поиска по любому сочетанию признаков практически малоприемлемым [22].

При анализе возможностей иерархической классификации для целей информационного поиска нельзя упускать из виду еще одно очень важное обстоятельство. В основе любой иерархической классификации, создаваемой по формально-логическим правилам, лежит отношение «род—вид». Это означает, что в принципе мы не имеем права использовать для деления объема понятия (класса) такие признаки, которые не дают нам видовых понятий (подклассов). Например, в идеальной иерархической классификации невозможно создать отдельный класс для документов о конструкциях или кузовах грузовых автомашин, так как понятие «конструкция» и «кузов» не являются видовыми по отношению к понятию «грузовые автомашины». Игнорирование этого условия приводит к тому, что оказываются нарушенными формально-логические правила построения иерархических классификаций со всеми вытекающими отсюда последствиями.

---

[22] J. Perry, A. Kent. Documentation and information retrieval. An introduction to basic principles and cost analysis. Cleveland, Ohio, Western Reserve University, 1957; стр. 4—13. J. R. Sharp. Some fundamentals of information retrieval. London. Deutsch, 1965, стр. 70—72.

Из сказанного выше следует, что применение иерархических классификаций для индексирования и поиска документов по сочетаниям характеристик, связанных друг другом не только родо-видовыми отношениями, наталкивается на определенные трудности.

Для преодоления этих трудностей при построении реальных иерархических ББК в качестве оснований деления используются не только видообразующие, но и другие признаки. Вот что пишет по этому поводу известный советский специалист по ББК З. Н. Амбарцумян:

«По нашему мнению, авторы излишне канонизируют применение отраслевого принципа в построении систематического каталога. Если внутри отраслевого отдела возникает возможность и практическая целесообразность введения подразделения по признаку «предмета», такое построение не будет противоречить основному принципу отраслевого построения каталога. В систематическом каталоге последовательность общих и частных делений определяется не только переходом от общего к частному, но и от целого к части» [23].

Отношение от целого к части — это разновидность ассоциативных отношений, а не отношение «род—вид». Независимо от того, насколько оправдано в практическом отношении это нарушение одного из важных принципов построения иерархических классификаций, оно лишает процесс создания таких классификаций объективной формально-логической основы и делает его полуинтуитивным процессом, зависящим от конкретных практических потребностей, от опыта составителей схемы и от многих других факторов. В результате этого структура иерархической классификации усложняется, в ней появляются пересекающиеся классы и т. д.

Ранее в качестве одного из важных достоинств иерархической ББК называлось то, что в ней классы располагаются один относительного другого в некотором полезном порядке. Но если иерархическая классификация используется только для образования классов, а эти классы в зависимости от изменяющихся практических потребностей будут по-разному располагаться один относительно другого, то такая классификация теряет это свое преимущество перед любой неиерархической системой.

В связи с этим необходимо отметить, что при подсчете числа иерархий, которые можно построить при $n=e=3$, было сделано предположение, что в практических условиях применима любая иерархия классов. Это предположение не всегда правильно, так как во многих случаях существует некоторая естественная иерархия классов, которая является безусловно более предпочтительной перед любыми другими. Например, для географической класси-

[23] З. Н. Амбарцумян. Некоторые спорные вопросы предметного каталога. «Сов. библиография», 1965, № 5, стр. 46.

фикации естественным будет следующий порядок классов: часть света—страна—республика (штат, земля)—область (край, автономная республика; воеводство; графство; департамент и др.)—район и т. д.

Очевидно, что если специфика классифицируемых предметов такова, что классы в иерархических системах располагаются в определенном естественном порядке, то такая классификация служит достаточно эффективным средством информационного поиска.

С другой стороны, имеется не меньше случаев, когда никакой естественной последовательности расположения классов в иерархической классификации не существует. Например, если бы нам пришлось создавать иерархическую схему классификации персонала по их физическим характеристикам (пол, возраст, рост, вес, цвет волос, цвет глаз и т. д.), то последовательность расположения классов в этой схеме могла бы быть самой различной. В таких случаях приемлемым является любой возможный вариант иерархии, причем число таких вариантов, как было показано выше, может быть очень большим. Кроме того, приемлемость всех возможных вариантов иерархии фактически лишает иерархическую классификацию такого важного свойства, как обозначение родовидовых отношений между классами. Поэтому применение иерархической классификации в таких условиях становится неоправданным.

Перейдем теперь к содержательному рассмотрению иерархической классификации как семантической системы, предназначенной для поиска документов. В основе любой иерархической классификации лежит представление о том, что вся совокупность накопленных человечеством знаний может быть разделена на взаимоисключающие классы и подклассы. В таких иерархических классификациях каждый класс делится лишь по одному основанию (признаку), давая одну серию подклассов. Классификационное дерево, получающееся в результате такого деления, является линейным и отличается большой жесткостью. Количество классов, подклассов и других подразделений в иерархической классификации ограничено, что равносильно ограниченности словаря, который может быть использован для целей индексирования. Проблему ограниченности такого словаря нельзя удовлетворительно решить путем бесконечного увеличения числа подразделений иерархической классификации, так как по мере его роста все более усложняется задача правильного отнесения того или иного документа к какому-то одному классу. Это объясняется многоаспектностью смыслового содержания почти любого документа, а также практической невозможностью обеспечить полную взаимоисключаемость классов в иерархической классификации. Поэтому систематизатор, несмотря на широкое применение перекрестных ссылок и дублирование, не может быть уверен в том, что приписанный им

документу поисковый образ (индекс или группа индексов) позволит отыскать этот документ при любом варианте запроса.

Развитие науки, как известно, характеризуется наличием двух противоположных тенденций: а) возрастающей дифференциацией, в результате которой каждая наука разделяется на все новые и новые ветви; б) увеличивающимся соприкосновением и взаимопроникновением не только смежных, но иногда очень далеких одна от другой наук, в результате чего появляются новые, ранее не существовавшие науки. Указанные тенденции развития науки, которые особенно ярко проявляются в наше время, отражают противоположные и вместе с тем тесно связанные между собой процессы анализа и синтеза, единство которых составляет одно из важнейших требований диалектического метода мышления.

Из сказанного выше следует, что любая претендующая на научность и перспективность ББК непременно должна учитывать процессы анализа и синтеза в развитии науки и иметь такую схему, которая бы позволяла адекватно отражать в классификации новые ветви уже сложившихся наук, новые науки и возникающие в результате дифференциации последних ветви новых наук.

Процесс дифференциации сложившихся отраслей науки и техники находит отражение в иерархической классификации, хотя такое отражение не всегда может быть признано удовлетворительным. Что же касается взаимопроникновения наук, образования новых научных дисциплин и их последующей дифференциации, то отражение этого процесса в иерархической классификации вообще невозможно. Действительно, иерархическая классификация не позволяет логически суммировать (объединять) принадлежащие к разным классам предметы и тем более разделять логические суммы таких предметов на иерархические подразделения. Например, в УДК невозможно логически суммировать такие предметы, как «цитология» и «генетика»; «химия», «геология» и «биология»; «математика» и «лингвистика» и т. д. Это является одним из основных недостатков любой иерархической классификации, ставящим ее в прямое противоречие с синтетической тенденцией развития современной науки. В любой иерархической классификации отдельные науки раз и навсегда разорваны разветвлениями жесткого классификационного дерева.

Особенно трудно включать в иерархическую классификацию новые межотраслевые предметы и области исследований, когда отсутствует даже общепринятая терминология для обозначения этих предметов. Между тем именно в начальные периоды исследовательских работ в новых областях информационные запросы по таким предметам наиболее многочисленны, и быстрое удовлетворение этих запросов имеет особенно большое значение.

Новые данные, получаемые в процессе научного познания объективного мира, требуют периодического изменения общей структуры той или иной реальной иерархической классификации.

Однако на разработку новых таблиц иерархических классификаций и особенно — на переиндексирование документов требуются очень большие затраты времени и труда. Это делает такую задачу практически неразрешимой, так как новый вариант иерархической классификации устаревает раньше, чем удается завершить работу по переиндексированию всех документов. Поэтому приходится идти по пути соответствующей модификации уже существующих таблиц иерархической классификации, что придает ей все более условный характер.

Приспособление устаревшей иерархической классификации — без изменения ее основной структуры — к непрерывно повышающемуся уровню наших знаний о природе и обществе — это очень трудная задача. На ее бесперспективность обращал внимание еще создатель первой десятичной классификации М. Дьюи. В одном из своих писем к С. Р. Ранганатану он писал:

«Как автор Десятичной классификации, которая в настоящее время получила более широкое распространение, чем все остальные [классификации] вместе взятые, я естественно уделяю особое внимание ее модификации, и я твердо убежден, что одна из самых серьезных ошибок состоит в том, что впустую расходуется время и деньги на «улучшение» какой-либо классификационной схемы» [24].

Ранее уже отмечалось, что многоаспектное индексирование документов, а также их поиск по любому сочетанию признаков возможны лишь тогда, когда иерархическая классификация содержит отдельные классы для любого сочетания признаков. Однако ни одна иерархическая классификация не может содержать такого множества классов. Справедливость такого утверждения можно показать следующим образом [25].

Допустим, что мы построили классификацию по основным цветам — красному, голубому и желтому — и решили использовать ее для группировки книг по цветам их переплетов. Если к нам поступит книга в голубом переплете, то задача отнесения ее к классу «голубые книги» решается легко. Однако совершенно иная ситуация возникнет тогда, когда нам попадется книга, например, в зеленом переплете. Зеленый цвет, как известно, образован путем смешения голубого цвета с желтым. Если переплет у книги светло-голубой, то на этом основании ее можно отнести к классу желтых книг. Но для этого необходимо точное цветовое восприятие и знание законов смешения цветов. Когда же переплет у книги зеленого цвета, то единственно возможным решением

---

[24] S. R. Ranganathan. Prolegomena to library classification. 2nd ed., London, The Library Association, 1957, стр. 420.

[25] Пример заимствован из кн.: W. F. Williams. Principles of automated information retrieval, Elmhurst, Ill., The Business Press, 1965, стр. 116.

будет дополнение нашей классификации еще одним классом для зеленых книг. Поскольку количество возможных цветов и их оттенков практически бесконечно, то бесконечен и процесс дополнения нашей классификации все новыми и новыми классами.

На этом примере ясно видна одна из основных проблем классификации — тенденция к бесконечному увеличению числа классов, требующихся для классифицирования предметов объективной действительности.

Другая трудность возникает тогда, когда нам попадает книга, переплет которой имеет, например, частично голубой и частично красный цвет. Очевидно, что мы отнесем эту книгу к двум классам — к классу голубых и к классу красных книг. Но при последующем поиске этой книги мы можем забыть, была ли данная книга красной И голубой или красной ИЛИ голубой.

Неприспособленность иерархической классификации для индексирования и поиска документов по сложным предметным комплексам можно показать также на следующем примере [26].

В последнем, 16-м издании Десятичной классификации Дьюи мы имеем следующую цепочку классов:

372.2 Организация начальной школы.
372.3 Изучение науки и природы.
372.4 Чтение и письмо.
372.5 Творческие науки (arts).
372.6 Языковые науки.
372.7 Арифметика.
372.8 Прочие вопросы (делить как 000—999).

Теперь предположим, что нам нужно заиндексировать документ об изучении природы в детских садах. Но предмет «детские сады» относится к классу 372.218, а предмет «изучение природы» — к классу 372.3. Отдельного класса для этого предметного комплекса в Десятичной классификации Дьюи нет. Поэтому однозначно заиндексировать документ по такому предмету или найти его в систематическом каталоге, организованном по Десятичной классификации Дьюи, очень трудно.

Еще один пример. Допустим, что нам необходимо заиндексировать по УДК метеорологический отчет о тайфуне «Нора», датированный июнем 1953 г. В этом отчете приводятся главным образом сведения о самом тайфуне, а также кратко сообщается об ущербе, который он причинил Гонконгу, Филиппинам, Тайваню и Японии. Отчет является чисто метеорологическим. Поэтому при хорошо разработанном разделе «Метеорология» в УДК он может быть заиндексирован, например, так: 551.515.23 «Нора» (914+512.317+521+522). Эти индексы позволят нам легко найти данный отчет, если нужно ответить на какой-либо «метеорологический» вопрос.

---

[26] D. J. Foskett. The London Education Classification. (Education Libraries Bull. Suppl. 6). London, University of London, 1963, стр. 2.

Но допустим, что поступил следующий «неожиданный» запрос: «В каких документах содержатся сведения об ущербе, который нанесли Японии различные бедствия в 1953 г.?» В этом случае очень мала вероятность того, что данный отчет вообще будет найден в систематическом каталоге, организованном по УДК, так как бедствия могут быть самого разнообразного характера (пожары, землетрясения, цунами, эпидемии, тайфуны, последствия испытания американских водородных и атомных бомб и т. д.), а отдельного класса для обобщенного понятия «бедствия» в УДК не предусмотрено [27].

Ранее подчеркивалось, что иерархические классификации обычно строятся на базе классификации наук с ее делением на отдельные отрасли. Но имеется множество предметов, особенно в области естествознания, медицины и техники, изучение которых не является задачей какой-либо одной науки. Например, одна и та же машина или аппарат может применяться в различных отраслях техники. Поэтому в иерархической классификации создается несколько классов для одного и того же предмета, изучаемого разными науками или рассматриваемого в разных аспектах. Документы о таком предмете попадают в разные классы, причем отсутствуют правила, позволяющие отыскать все эти классы при информационном поиске. Это явление в определенном смысле аналогично явлению лексической синонимии в естественных языках. Оно сильно затрудняет (а иногда делает почти невозможным) отыскание в систематическом каталоге всех документов по любому межотраслевому предмету.

Описанное выше явление «синонимии» можно продемонстрировать на примере УДК. В этой классификации предмет «смазочные вещества» может быть отнесен к классам 62-72 и 621.89. Вполне вероятно его отнесение также к классу 665. Предмет «электроизмерительные приборы» в УДК может быть отнесен к классам 537.74 и 621.317.7; предмет «гибридизация» — к классам 575.12 и 631.523; предмет «наследственность» — к классам 575.1 и 581.169 и т. д.

Итак, невозможность иметь в иерархической классификации отдельные исчерпывающие классы для любого предмета и предметного комплекса делает такую классификацию малопригодной для информационного поиска по любым, заранее не предусмотренным сочетаниям предметов, а также по единичным предметам и предметам межотраслевого характера.

Теперь можно попытаться суммировать основные недостатки иерархических классификаций как семантической системы, пред-

---

[27] Пример заимствован из статьи: J. A. Schüller. Experience with indexing and retrieving by UDC and Uniterms. «Aslib Proceedings», 1960, v. 12, № 11, стр. 388.

назначенной для поиска документов. Эти недостатки состоят в следующем:

1. Невозможность многоаспектного индексирования документов.

2. Непригодность для информационного поиска по любому сочетанию признаков, а также по узкопредметным и межотраслевым запросам.

3. Недостаточная глубина деления, из-за чего в классификации могут отсутствовать классы для специфических предметов.

4. Невозможность обеспечения взаимоисключаемости классов (явление «синонимии»).

5. Недостаточная эффективность механизма пересмотра классификации. Это означает, что в ней отсутствуют подразделения для новых предметов до тех пор, пока не будет опубликовано новое издание таблиц этой классификации. Таким образом, иерархическая классификация всегда отстает от достигнутого уровня развития науки и техники.

Строго говоря, непригодность иерархических классификаций для поиска документов по единичным и межотраслевым предметам нельзя считать их недостатком. Таково одно из их внутренних свойств, которое необходимо принимать per se.

Наряду с существенными недостатками, иерархические классификации имеют и важные достоинства, благодаря которым различные классификации в течение многих веков используются в качестве основного средства для тематического поиска литературы в библиотеках. Ниже перечислены основные достоинства иерархических классификаций:

1. Классификация — один из привычных методов, которыми пользуются люди для определения предметов. В. И. Ленин писал по этому поводу: «Что значит дать „определение"? Это значит, прежде всего, подвести данное понятие под другое, более широкое» [28].

2. Иерархическая классификация пригодна для каталогов и указателей, имеющих любую физическую форму.

3. Для составления и использования систематического каталога или указателя не требуется никаких специальных средств и устройств.

4. На индексирование документов по иерархической классификации, как правило, расходуется меньше квалифицированного труда, чем на индексирование по фасетным и алфавитно-предметным классификациям.

5. При информационном поиске по тематическим запросам ручной систематический каталог или указатель едва ли уступает механизированным и автоматизированным системам по эффек-

---

[28] В. И. Ленин. Полное собрание сочинений, т. 18, стр. 149.

тивности и быстродействию и превосходит их по дешевизне поиска.

6. Для шифровки классов в иерархической классификации обычно применяются арабские цифры и буквы латинского алфавита, имеющие широкое международное распространение. Это превращает ту или иную иерархическую классификацию в специализированный международный язык, понятный специалистам разных стран.

Подводя итог общему анализу иерархической классификации как семантической системы, предназначенной для поиска документов, можно прибегнуть к аналогии и сравнить иерархическую классификацию с двуязычным разговорником, содержащим ограниченное число готовых фраз. A priori очевидно, что в таком разговорнике не могут содержаться все фразы, необходимые нам для выражения любой мысли. Если же попытаться составить разговорник, который бы содержал все возможные высказывания на данных двух языках, то такой гипотетический разговорник имел бы настолько большой объем, что пользоваться им было бы практически невозможно.

С другой стороны, при правильном использовании двуязычные разговорники вполне успешно выполняют свое назначение. Точно так же иерархические классификации при определенных условиях являются весьма ценным инструментом для тематического поиска документов.

## ФАСЕТНЫЕ БИБЛИОТЕЧНО-БИБЛИОГРАФИЧЕСКИЕ КЛАССИФИКАЦИИ

В последние годы все большее распространение, особенно в Великобритании, получают так называемые *аналитико-синтетические*, или *фасетные классификации*. В основе этих классификаций лежит *фасетный анализ*, идея которого принадлежит выдающемуся индийскому библиографу и библиотековеду Ш. Р. Ранганатану.

Сущность фасетного анализа состоит в следующем. Сначала тщательно анализируется отрасль науки или техники, для которой составляется фасетная классификация, а также изучается некоторый массив документов по этой отрасли. На основе такого анализа составляется перечень основных категорий предметов, применяемых в данной отрасли. Эти категории называются *фасетами*. Кроме того, из изучаемых документов выписываются все существенные термины, относящиеся к данной отрасли науки или техники. Эти термины группируются по фасетам, т. е. объединяются в соответствующие классы. Множество терминов, обозначающих различные материалы, образует фасет «Материалы»; множество терминов, обозначающих различные процессы, — фасет «Процессы» и т. д. Каждый термин фасета называется *фокусом*.

Фасетный анализ — весьма мощный прием, который может быть успешно применен для разных целей, в том числе и для построения дескрипторных ИПЯ. В частности, применение фасетного анализа при построении иерархических ББК позволяет предотвратить образование пересекающихся классов на одном уровне иерархии.

Следующие этапы построения фасетной классификации — деление фасетов на субфасеты, субсубфасеты и т. д. (если необходима бóльшая детализация); установление полезной, но фиксированной последовательности расположения терминов внутри фасетов и фасетов — в схеме классификации; обозначение соответствующими шифрами фасетов и терминов, входящих в эти фасеты; установление фиксированной последовательности расположения фасетов в поисковом образе (эта фиксированная последовательность называется *фасетной формулой*).

Процедура индексирования документов по фасетной классификации осуществляется следующим образом. Сначала на естественном языке формулируется (мысленно или на бумаге) основное смысловое содержание документа. Затем этот смысл выражается в терминах фасетной классификации, т. е. при помощи цепочки фокусов, взятых из фасетов и расположенных в фиксированном порядке. Часто вместо фокусов используются их шифры. Такая процедура позволяет создавать классы для документов, предметы которых выражаются сочетанием нескольких разноаспектных характеристик.

Из сказанного выше следует, что между фасетными и иерархическими ББК имеется весьма существенное различие. Составители иерархических классификаций стремятся дать в каком-то систематическом порядке перечень всех предметных комплексов, которые уже описаны в литературе. Но даже для сравнительно узкой отрасли науки или техники невозможно перечислить все такие предметные комплексы, так как их число огромно. Кроме того, в результате прогресса науки и техники непрерывно появляются новые предметные комплексы. Поэтому перечислительные классификации в принципе не могут быть исчерпывающими.

При построении фасетной классификации не делается попытки перечислить в ее таблицах все предметные комплексы. Вместо этого дается некоторое множество элементарных термов («строительных кирпичей»), из которых можно построить огромное число наименований предметных комплексов.

Основные принципы построения и использования фасетных классификаций можно проиллюстрировать на примере, относящемся к тяжелому машиностроению [29]. Пусть даны фрагменты фасетов «Общие свойства», «Материалы», «Типы разрушений» и «Напряжения и нагрузки».

[29] См.: F. W. Lancaster. Engineering information storage: indexing vs. classification. «Machine Design», 1965, v. 37, № 1, стр. 106.

| | | | |
|---|---|---|---|
| *A* | Общие свойства | *B* | Материалы |
| *Ab* | Конфигурация | *Bc* | Металлы |
| *Ac* | Трубчатая | *Bcc* | Черные металлы |
| | | *Bcd* | Стали |
| | | *Bcf* | Хромистые стали |
| | | *Bcfi* | Хромоникелевые стали |
| | | *Bqt* | Никелевые сплавы |
| *K* | Типы разрушений | *L* | Напряжения и нагрузки |
| *Kg* | Ползучесть | *Lb* | Растяжение |
| *Kgb* | Длительная прочность | *Ld* | Кручение |
| *Ki* | Хрупкий излом | *Le* | Срез |
| *KL* | Продольный изгиб | *Lg* | Сжатие |
| *Kp* | Усталость | *Lm* | Удар |
| *Kpf* | Коррозионная усталость | | |

Из терминов, перечисленных в этих фасетах, можно построить классы для следующих сложных понятий:

*KpBcfi*=усталость хромоникелевых сталей
*LmKpBcf*=усталость хромистых сталей при ударной нагрузке
*KgbBqtAc*=длительная прочность трубчатых элементов из никелевых сплавов и т. д.

Важным недостатком перечислительных классификаций является то, что они не позволяют строить классы путем применения любого сочетания характеристик. Если в иерархической классификации характеристика была применена на какой-то ступени деления объема понятия, то такая характеристика уже не может быть исключена из содержания видовых понятий, образующихся в результате применения других характеристик на последующих ступенях деления. Например, в приведенной на рис. 42 перечислительной классификации видовой признак «2» не может быть опущен в индексе *B2c*.

Фасетная классификация не связывает нас необходимостью пользоваться только готовыми классами. При использовании фасетной классификации мы можем строить названия классов (а следовательно, и сами классы) из разных сочетаний фокусов, пропуская в фасетной формуле ненужные нам фасеты. Если в иерархической классификации вообще невозможно, не нарушая всей структуры классификационного «дерева», построить класс *Bc;* то в фасетной классификации одинаково возможен как класс *B2c*, так и класс *Bc*.

Итак, по сравнению с классификациями перечислительного типа фасетные классификации значительно облегчают многоаспектное индексирование документов. Но и такие классификации фактически не обеспечивают информационного поиска по любому сочетанию терминов (признаков), так как для этого в систематическом каталоге, организованном по фасетной классификации, пришлось бы иметь отделы для всех возможных сочетаний и перестановок терминов, взятых из разных фасетов. Число таких

отделов было бы столь же велико, как и число всех возможных классов и подклассов в иерархической классификации. К такому выводу можно прийти даже путем следующего простого рассуждения.

Допустим, что у нас имеются фасеты:

«Типы гражданских самолетов»
- *Ac* Пассажирские самолеты
- *Ad* Транспортные самолеты
- *Ae* Специальные самолеты

«Число моторов»
- *Cd* Одномоторные
- *Cf* Двухмоторные
- *Ch* Четырехмоторные

«Расположение крыла»
- *Ec* Высокое расположение крыла
- *Ee* Среднее расположение крыла
- *Eg* Низкое расположение крыла

Нетрудно понять, что если расположить термины первого фасета на одной горизонтали, а затем приписать к каждому из этих терминов поочередно все термины второго фасета и после этого повторить описанную процедуру, используя термины третьего фасета, то мы получим иерархическую классификацию, приведенную на рис. 42. Таким образом, индуктивное построение иерархической классификации (при помощи фасетного анализа) приводит к такому же конечному результату, что и ее дедуктивное построение (путем последовательного деления классов) при условии, что соблюдаются основные формально-логические правила.

Из сказанного выше следует, что число всех возможных классов фасетной классификации во всяком случае не меньше числа иерархий, которые можно построить для эквивалентной ей иерархической классификации [30].

Однако даже общее сопоставление иерархической и фасетной классификаций показывает, что фасетные классификации обладают рядом существенных преимуществ перед иерархическими классификациями. Основные из этих преимуществ состоят в следующем:

1. Фасетные классификации позволяют производить многоаспектное индексирование документов.

2. Фасетные классификации дают возможность удобно собирать в одном месте все аспекты какого-либо предмета или темы. В общих иерархических схемах эти аспекты могут быть рассеяны по разным подразделениям (явление «синонимии»), причем отсутствуют правила отыскания таких подразделений.

3. Фасетные классификации допускают большую глубину деления понятий.

[30] Дж. Шарп утверждает, что для фасетной классификации число $N$ даже больше, чем для эквивалентной ей иерархической классификации. См.: J. R. Sharp. Указ. соч., стр. 72—73. Однако с этим утверждением нельзя согласиться.

Однако нет никаких существенных причин, которые бы запрещали применение характеристик из вспомогательных классов (фасеты 1—5) в качестве непосредственных входов в организованный по УДК систематический указатель (каталог). В частности, это вполне возможно при реализации такого указателя (каталога) с помощью ЭЦВМ. Очевидно, что применение характеристик из вспомогательных классов в качестве непосредственных входов в указатель (каталог) существенно повысило бы его эффективность как документальной ИПС.

До сих пор речь шла о «словарном составе» УДК, т. е. об именах (или условных обозначениях) основных и вспомогательных классов. Было показано, что эти имена (или условные обозначения) могут выступать в функции дескрипторов. Иначе обстоит дело с правилами построения, представляющими «синтаксис» УДК. Исчерпывающий критический анализ этих правил можно найти в работе Ж. Перро [33].

Основными недостатками правил построения в УДК являются:

1) недостаточная специфичность правил построения (небольшое число символов используется для выражения значительно большего числа различных связей между характеристиками);

2) правила построения не позволяют выражать в линейном виде многомерные связи между характеристиками, что приводит к появлению многозначных выражений (индексов);

3) неоднозначность правил построения (не нарушая этих правил, два систематизатора могут по-разному выразить одно и то же отношение между характеристиками).

Эти недостатки ныне действующих в УДК правил построения весьма затрудняют их формальное применение, т. е. использование в механизированных ИПС.

## НОТАЦИЯ И АЛФАВИТНЫЕ УКАЗАТЕЛИ К КЛАССИФИКАЦИОННЫМ СХЕМАМ

Основная таблица ББК — как иерархической, так и фасетной — должна быть снабжена двумя важными вспомогательными средствами:

а) нотацией, или системой шифров (индексов),

б) алфавитно-предметным указателем.

Назначение шифров (индексов), которые присваиваются каждому подразделению ББК, — облегчить расстановку классифицируемых документов (или их описаний) в ИПС и их последующий поиск, сделать эту работу чисто механической. Для этого названиям подразделений ББК, расположенных в нужном порядке, приводятся в соответствие определенные символы (обычно в качестве таких символов используются цифры или буквы алфавита).

[33] J. Perreault. Categories and relators: a new schema. «Revue internationale de la documentation», 1967, v. 32, № 4, стр. 136—144.

Эти символы (шифры, индексы) проставляются на каждом классифицируемом документе (или его описании) и точно указывают место, в которое должен быть помещен этот документ (или его описание). Если бы в библиотеке, применяющей систематическую расстановку книг в хранилище, на книгах не были проставлены их классификационные шифры, то книгу после каждой ее выдачи нужно было бы классифицировать вновь, чтобы найти в книгохранилище то место, где она должна стоять.

Практика показывает, что предпочтение следует отдать иерархическим системам нотации, полностью отражающим родо-видовые отношения между классами. Однако используются также и порядковые (неиерархические) системы нотации.

Проблема нотации в ББК очень сложна, и ее подробный анализ выходит за рамки задачи, поставленной перед данной работой. Здесь необходимо лишь подчеркнуть, что нотация, будучи производной от классификационной схемы, никоим образом не определяет эффективности ББК как средства поиска документов, хотя и может оказывать на нее определенное влияние.

Алфавитно-предметный указатель к ББК (или к систематическому каталогу) предназначен для следующих целей [34]:

а) определения по названиям предметов их местонахождения в ББК (или в систематическом каталоге);

б) показа всех аспектов предмета, которые могут быть рассеяны в ББК (или в систематическом каталоге). Например:

| | | |
|---|---|---|
| Болезни: | Животные | 636.089 |
| Болезни: | Медицина | 616 |
| Болезни: | Растения: Сельское хозяйство | 632 |
| Болезни: | Пшеница: Сельское хозяйство | 633.11-2; |

в) перевода названий предметов с естественного языка на язык ББК (выполнение функции исчерпывающего словаря). При этом может оказаться, что в ББК отсутствует класс, обозначенный названием данного предмета. Тогда алфавитно-предметный указатель отсылает потребителя к вышестоящему классу, непосредственно включающему этот предмет. Кроме того, алфавитно-предметный указатель содержит все синонимы названий предметов, причем в таком указателе синониму обычно приведен в соответствие шифр класса, к которому этот синоним принадлежит (отсылки «см.» употребляются редко).

Алфавитно-предметный указатель к ББК (или к систематическому каталогу) может быть разной степени сложности: от простейшего, когда в указателе приводятся названия лишь основных классов ББК, до очень сложного — «цепного» указателя. Именно этот тип алфавитно-предметного указателя к ББК (или к система-

[34] J. Mills. The Universal Decimal Classification. (Rutgers series on systems for the intellectual organization of information, v. 1). New Brunswick, N. J., The Rutgers University Press, 1964, стр. 58—59.

тическому каталогу) представляет наибольший интерес для информационного поиска. Однако методика составления «цепного» указателя, а также его основные достоинства и недостатки будут рассмотрены в следующем разделе, посвященном алфавитно-предметным системам.

## АЛФАВИТНО-ПРЕДМЕТНЫЕ КЛАССИФИКАЦИИ

Алфавитно-предметная классификация представляет собой ИПЯ, основной словарный состав которого состоит из упорядоченного по алфавиту множества слов, словосочетаний и фраз естественного языка, обозначающих предметы какой-либо отрасли науки или практической деятельности. Е. И. Шамурин дает следующее определение алфавитно-предметной классификации: «Предметная классификация — классификация, в которой каждому предмету или вопросу отводится только один индекс, собирающий всю литературу в нем, независимо от разреза (аспекта), в котором предмет или вопрос в произведениях печати рассматривался» [35].

Для рассмотрения структуры и основных свойств алфавитно-предметных классификаций введем несколько дополнительных определений [36].

*Предметный заголовок* — это слово, словосочетание или фраза естественного языка, используемая для обозначения предмета всего документа, части документа (текста) или информационного запроса.

*Предметный словник* — это упорядоченное по алфавиту множество предметных заголовков, используемых для построения какого-либо алфавитно-предметного каталога или указателя.

*Предметная рубрика* — это совокупность предметного заголовка с описанием или адресом хранения хотя бы одного документа или с указанием тех мест в тексте документа, основная тема которых обозначается этим предметным заголовком.

В алфавитно-предметном указателе к систематическому каталогу предметную рубрику образует совокупность предметного заголовка с шифром класса, к которому относятся документы по предмету, обозначенному этим заголовком.

В одном из своих недавних выступлений Ч. Л. Берньер [37] так определил сущность алфавитно-предметной классификации (индексирования):

---

[35] Е. И. Шамурин. Словарь книговедческих терминов. М., «Сов. Россия», 1958, стр. 229.

[36] Необходимо особо подчеркнуть, что в задачу данного раздела не входит даже краткое изложение столь сложной темы, как теория и методика алфавитно-предметной классификации.

[37] Ч. Л. Берньер — известный американский специалист по информатике, который в течение многих лет руководил составлением указателей к РЖ «Chemical Abstracts».

«. . . предмет — это то, чему посвящен разговор, дискуссия или статья. Предметный заголовок — это термин, который выбран для того, чтобы направить или привести потребителя (inquirer) к предмету, а предметное индексирование — это метод указания предмета потребителю. Предметизатор должен выявить один или более предметов, о которых идет речь, подобрать один или более терминов для обозначения этих предметов, после чего заменить данные термины стандартными, применяемыми при индксировании» [38].

Алфавитно-предметные классификации предназначены для построения указателей и каталогов, используемых главным образом для узкопредметного поиска. В таких указателях и каталогах под предметным заголовком даются сведения (адресные шифры или библиографические описания) по возможности о всех документах, предмет которых обозначен данным заголовком. «В каждой предметной рубрике, — пишет С. Г. Карпова, — непосредственно под точным названием данного предмета (вопроса), должна собираться вся основная литература о нем, то есть все произведения, основной темой которых является данный предмет (вопрос)» [39].

Алфавитно-предметные классификации могут быть двух типов:

а) алфавитно-систематические;

б) словарные.

В алфавитно-систематической классификации предметный словник состоит из названий предметов, которые подчинены названиям классов таких предметов. Число уровней иерархии в таком предметном словнике определяется практическими требованиями, предъявляемыми непосредственно к данной классификации. Порядок расположения предметных заголовков в таком словнике — алфавитный. Ниже приводится пример построения предметного заголовка в алфавитно-систематической классификации:

ЭЛЕМЕНТАРНЫЕ ЧАСТИЦЫ — БАРИОНЫ — НУКЛОНЫ — НЕЙТРОНЫ

В словарной классификации предметный словник состоит только из расположенных в алфавитном порядке названий самих предметов. Идеальной моделью такой классификации можно считать именной указатель или каталог. В словарной классификации приведенный выше предметный заголовок будет иметь следующий вид:

НЕЙТРОНЫ

---

[38] Цит. по: J. Metcalf. Alphabetical subject indication of information. (Rutgers series on systems for the intellectual organization of information, v. 3). New Brunswick, N. J., The Rutgers University Press, 1965, стр. 141.

[39] С. Г. Карпова. Методика предметизации технической литературы. Л., Гос. Публ. б-ка им. М. Е. Салтыкова-Щедрина, 1957, стр. 7.

В настоящее время алфавитно-систематические классификации применяются сравнительно редко. Поэтому в данной работе они не рассматриваются, и основное внимание уделяется анализу возможностей алфавитно-предметных классификаций словарного типа.

В алфавитно-предметной классификации основными структурными элементами словаря считают предметный заголовок, предметный подзаголовок и предметную запись [40].

*Предметный подзаголовок* — это слово, словосочетание или фраза естественного языка, которая обозначает аспект рассмотрения предмета, указанного в предметном заголовке. Подзаголовок может также обозначать подкласс предметов, которые входят в более широкий класс, обозначенный предметным заголовком. (Подзаголовок первой ступени можно считать заголовком по отношению к подзаголовку второй ступени и т. д.). Подзаголовки обычно включаются в предметный словник. Порядок расположения подзаголовков в пределах заголовка — алфавитный.

*Предметная запись* — это слово, словосочетание или фраза естественного языка, уточняющая и конкретизирующая подзаголовок. Как и подзаголовок, она может также обозначать в пределах заголовка подкласс предметов третьего уровня иерархии. Предметные записи обычно не включаются в словник. Иногда в качестве предметных записей используются заглавия документов или их библиографические описания. Последовательность расположения предметных записей в пределах подзаголовка может быть как алфавитной, так и любой другой.

Такое деление структурных элементов словарного состава алфавитно-предметных классификаций на заголовки, подзаголовки и предметные записи весьма условно. Действительно, если подзаголовок обозначает нулевой (пустой) класс, то такой подзаголовок может быть опущен. Наоборот, если в обозначаемом им классе содержится слишком много членов, то этот подзаголовок целесообразно сделать заголовком. Не во всех алфавитно-предметных классификациях применяются предметные записи. Поэтому в данной работе различие между заголовками, подзаголовками и предметными записями не проводится, и структурной единицей алфа-

---

[40] Н. М. Сагалович. Некоторые вопросы теории и практики построения предметного указателя. Сб. «Вопросы механизации и автоматизации информационных работ». М., ВИНИТИ, 1962, стр. 16—31.

Л. В. Зильберминц. Предметные указатели к реферативным журналам и их роль в связи с развитием информационно-поисковых систем. «Труды Ленингр. ин-та культуры им. Н. К. Крупской», 1964, т. 15, стр. 195—208; А. В. Клёнов. Библиотечные каталоги. М., Всесоюз. кн. палата, 1963, стр. 289—309; Recommendations for the preparation of indexes for books, periodicals and other publications. British Standard 3700 : 1964, L., British Standards Institution, 1964.

витно-предметной классификации считается предметный заголовок, который может состоять из нескольких сегментов, а каждый сегмент — из одного и более слов (терминов). Например, структуру

АВТОМОБИЛИ ГРУЗОВЫЕ, ТРЕХОСНЫЕ
— ПРОИЗВОДСТВО
— СССР

мы рассматриваем как сложный предметный заголовок

АВТОМОБИЛИ ГРУЗОВЫЕ, ТРЕХОСНЫЕ—ПРОИЗВОДСТВО—СССР

Впервые система основных правил составления и применения алфавитно-предметных классификаций была разработана в наиболее полном виде известным американским библиографом Ч. А. Кеттером в его книге «Rules for a dictionary catalogue», первое издание которой [41] вышло в 1876 г. Именно в США алфавитно-предметные классификации получили наиболее широкое применение. В последние годы значительный вклад в теорию и практику алфавитно-предметных классификаций был внесен английским библиографом Коутсом [42].

Одно из важнейших требований, предъявляемых к алфавитно-предметной классификации, состоит в том, чтобы каждый заголовок имел точно такой же объем (экстенсионал), что и предмет, для обозначения которого этот заголовок предназначен. Объем понятия — это множество предметов, каждый из которых обладает всеми признаками, составляющими содержание данного понятия. Из этого определения следует, что подстановка названия каждого предмета, входящего в объем понятия, на место $x$ в формуле одноместного предиката $P(x)$ всегда дает истинные суждения. Данное требование должно в максимальной степени выполняться при формулировании предметных заголовков.

Другое важное требование заключается в том, чтобы обеспечить формулирование предметов документов и поисковых предписаний в одних и тех же терминах, т. е. единообразное употребление предметных заголовков не только предметизатором, но и потребителями информации, осуществляющими информационный поиск. Очевидно, что чем выше степень формализации предметного словника, тем больше такое единообразие и тем эффективнее действует алфавитно-предметный каталог или указатель.

При составлении предметного словника необходимо решить три задачи:

---

[41] C. A. Cutter. Rules for a dictionary catalogue. 4th ed. Washington, U. S. Government Printing Office, 1904.

[42] E. J. Coates. Subject catalogues: heading and structure. London, The Library Association, 1960.

а) устранить синонимию слов, словосочетаний и фраз, используемых в качестве предметных заголовков;

б) выделить основные, ведущие слова в словосочетаниях, используемых в качестве предметных заголовков;

в) обозначить смысловые (парадигматические) связи между названиями предметов, разорванные алфавитным расположением предметных заголовков.

Синонимия предметных заголовков в словнике устраняется следующим способом. Один из синонимов назначается предметным заголовком, а остальные снабжаются отсылками «см.» к этому предметному заголовку.

Правила выделения ведущих слов в словосочетаниях значительно сложнее. При решении такой задачи нужно иметь в виду, что ведущее слово предметного заголовка вместе с поясняющими его словами выполняет функцию аннотации, на основании которой потребитель решает вопрос о соответствии или несоответствии его информационному запросу документов, перечисленных под этим заголовком. Кроме того, предметный заголовок должен иметь такую форму, которая обеспечивает быстрое отыскание этого заголовка в каталоге или указателе, а также группировку в одном месте каталога или указателя близких по смыслу предметов, т. е. «комплексирование». Это способствует сокращению времени информационного поиска в каталоге или указателе. В связи с этим следует отметить, что формализация предметного словника — основное средство сокращения времени информационного поиска в алфавитно-предметном каталоге или указателе. Принципы и правила формализации предметного словника должны быть хорошо известны не только предметизатору, но и потребителям информации.

Для обеспечения правильности использования предметного словника теми потребителями, которым не известны принципы и правила его построения, этому словнику придается определенная избыточность. Эта избыточность создается за счет введения в предметный словник не только синонимов, но и всех возможных структурных вариантов заголовков. Эти структурные варианты снабжаются отсылками «см.», которые направляют потребителя к той форме словосочетания или фразы, которая принята для предметного заголовка. Например:

а) ГАЗОВОЕ ХРОМИРОВАНИЕ
см.: ХРОМИРОВАНИЕ ГАЗОВОЕ
ЗУБЧАТЫЕ НАСОСЫ
см.: НАСОСЫ ЗУБЧАТЫЕ

б) ТУРБИННОЕ БУРЕНИЕ НА НЕФТЬ
см.: НЕФТЯНЫЕ СКВАЖИНЫ — БУРЕНИЕ ТУРБИННОЕ

Однако с увеличением избыточности предметного словника растет его объем, что является весьма нежелательным.

Смысловые (парадигматические) связи между названиями предметов в словнике обозначаются при помощи ссылок «см.» и «см. также». Ссылками «см.» обозначаются отношения «вид—род». Например:

БАЗАЛЬТЫ КАК ЦЕМЕНТНОЕ СЫРЬЕ
см. ЦЕМЕНТНОЕ СЫРЬЕ

Ссылки «см. также» применяются для обозначения отношений «род—вид», а также ассоциативных связей типа «часть—целое», «процесс—орудие», «процесс—материал», «материал—изделие, продукт», «процесс—объект» и т. д. Например[43]:

а) ПНЕВМАТИЧЕСКИЕ МАШИНЫ
см. также: КОМПРЕССОРЫ
ТУРБИНЫ
б) РУДНИЧНЫЙ ВОДООТЛИВ — МЕХАНИЗАЦИЯ
см. также: РУДНИЧНЫЕ НАСОСЫ
в) КОЖИ — КРАШЕНИЕ
см. также: КРАСИТЕЛИ ДЛЯ КОЖИ
г) ДОЛОМИТ — ПЕРЕРАБОТКА
см. также: КСИЛОЛИТ — ПРОИЗВОДСТВО ИЗ ДОЛОМИТА
д) ДУБЛЕНИЕ
см. также: МЕХА — ДУБЛЕНИЕ
е) КОЛЕНЧАТЫЕ ВАЛЫ
см. также: ТРАКТОРНЫЕ ДВИГАТЕЛИ — КОЛЕНЧАТЫЕ ВАЛЫ

Необходимо отметить, что последовательное применение ссылок «см.» и «см. также» для обозначения родо-видовых связей между названием предметов неминуемо приведет к воспроизведению в предметном словнике той или иной иерархической классификации. Это может быть предотвращено путем введения определенных ограничений в использовании ссылок «см.» и «см. также» для обозначения родо-видовых связей.

Алфавитно-предметная классификация позволяет вводить новые предметные заголовки, когда в этом возникает необходимость. Но она требует строгого контроля за словами и словосочетаниями, используемыми в качестве предметных заголовков, а также за системой смысловых связей между такими заголовками.

В алфавитно-предметной классификации используются следующие типы заголовков:

---

[43] Все примеры здесь и далее заимствованы из книги С. Г. Карповой. Методика предметизации технической литературы. Л., Гос. Публ. б-ка им. М. Е. Салтыкова-Щедрина, 1957.

1. Простые существительные, причем некоторые из них снабжаются пометой, устраняющей их многозначность. Например:

КРИСТАЛЛЫ
РЕШЕТКИ (*кристаллов*) и т. д.

2. Двучленные и многочленные именные словосочетания [44], которые могут быть следующих видов:

а) словосочетания вида $AS$ и $AAS$, где $A$ — прилагательное, а $S$ — существительное. Например:

ЛИТЕЙНОЕ ПРОИЗВОДСТВО
ЖАРОПРОЧНЫЕ АВИАЦИОННЫЕ СПЛАВЫ

б) беспредложные словосочетания вида $SS$, $S\,(AS)$, $(AS)\,S$, $(AS)\,(AS)$, $(AS)\,(AAS)$ и т. д. Например:

ЗАКАЛКА МЕТАЛЛОВ
ПОВРЕЖДЕНИЕ ПАРОВЫХ КОТЛОВ
УЛЬТРАЗВУКОВАЯ ОБРАБОТКА МАТЕРИАЛОВ и т. д.

В словосочетаниях вида (а) и (б) слова связаны атрибутивными, объектными и обстоятельственными отношениями;

в) предложные словосочетания вида $SpS$, $SpAS$, $(AS)pS$, $(AS)p(AS)$ и т. д., где $p$ — предлог «в», «из», «для» или «как», выражающий соответствующие обстоятельственные отношения (пространственные; целевые; образа действия; превращение чего-л. во что-л.; свойства чего-л. по составу, материалу; в функции чего-либо). Например:

ЭЛЕКТРОСВАРКА В СУДОСТРОЕНИИ
КОМПРЕССОРЫ В АЗОТНОМ ПРОИЗВОДСТВЕ
АБРАЗИВНЫЕ МАТЕРИАЛЫ ДЛЯ СТЕКЛА
ЭКСКАВАТОРНЫЕ РАБОТЫ В ДОРОЖНОМ СТОИТЕЛЬСТВЕ

3. Грамматически не связанные комбинации слов типа $S—S$, $S—AS$, $AS—S$, $AS—AS$, $S—SpS$, $S—Sp(AS)$, $AS—SpS$, $S—S—S$, $AS—S—S$, $S\&S$, $AS—S\&S$ и т. д. Например:

СВАИ — ЗАБИВКА
СТЕКЛО — ЭЛЕКТРИЧЕСКИЕ СВОЙСТВА
КЕРАМИЧЕСКИЕ ЗАВОДЫ — ОБОРУДОВАНИЕ
ЦЕМЕНТООБЖИГАТЕЛЬНЫЕ ПЕЧИ — ОХЛАЖДЕНИЕ ВОДОЙ
КСИЛОЛИТ — ПРОИЗВОДСТВО ИЗ ДОЛОМИТА
ОЛИФЫ — ПРОИЗВОДСТВО ИЗ РЫБЬЕГО ЖИРА
ДОМЕННЫЙ ШЛАК — ИСПОЛЬЗОВАНИЕ В СТРОИТЕЛЬСТВЕ

---

[44] Именными называются словосочетания, в которых существительное выступает в роли главного слова.

ГОРЮЧИЕ СЛАНЦЫ — СЖИГАНИЕ В ПАРОВОЗНЫХ ТОПКАХ
НЕФТЬ — ДОБЫЧА — ГИДРОМЕХАНИЗАЦИЯ
ТКАЦКИЕ СТАНКИ — ДЕТАЛИ — ПРОИЗВОДСТВО
ЛИТЕЙНОЕ ПРОИЗВОДСТВО — ИЗОБРЕТАТЕЛЬСТВО И РАЦИОНАЛИЗАЦИЯ и т. д.

Теперь, когда в самом общем виде изложены основные сведения о структуре и принципах построения алфавитно-предметных классификаций, можно сделать попытку оценить эффективность таких классификаций при их использовании для поиска документов по различным сочетаниям характеристик (предметов). Предположим, что у нас имеется алфавитно-предметный указатель (каталог), который в дальнейшем будем обозначать аббревиатурой АПУ. Пусть в этот указатель введена информация о документе $d_i$. Основная тема этого документа выражена сложным предметным заголовком $ABCDE$, где $A$, $B$, $C$, $D$ и $E$ — отдельные слова (термины).

Для отыскания документа $d_i$ в АПУ необходимо соответствующим образом сформулировать поисковое предписание, причем это должно быть сделано в терминах данной алфавитно-предметной классификации. Возможны три основных варианта поискового предписания:

а) назван лишь один из терминов, входящих в сочетание $ABCDE$;

б) названо все сочетание $ABCDE$ (последовательность расположения терминов в этом сочетании фиксирована);

в) названы не все, а лишь несколько терминов, входящих в сочетание $ABCDE$.

Если АПУ предназначен для информационного поиска по предписаниям типа (а), то сведения о документе $d_i$ должны быть приведены в этом указателе под каждым термином, входящим в заголовок $ABCDE$. Такой АПУ будет содержать столько вариантов предметных заголовков, сколько терминов используется в нем для индексирования.

Для информационного поиска по предписаниям типа (б) сведения о документах по какому-либо предмету должны быть приведены в АПУ под одним заголовком, и это очень удобно. Однако весьма мала вероятность того, что формулируемые потребителями поисковые предписания будут иметь такое же содержание и структуру, что и поисковые образы (сложные предметные заголовки), присвоенные документам при индексировании. Поэтому АПУ, предназначенные для информационного поиска по предписаниям типа (б), применяются очень редко.

Одним из немногих примеров таких АПУ является формульный указатель к РЖ «Chemical Abstracts». В этом указателе приводятся эмпирические формулы описанных в РЖ химических соединений.

При поиске публикаций о каком-либо химическом соединении требуется, чтобы заданная эмпирическая формула искомого соединения точно совпала с формулой, приведенной в указателе.

Подавляющее большинство реальных информационных запросов требует проведения информационного поиска по предписаниям типа (в). Но в этом случае очень трудно обеспечить достаточно быстрое отыскание в АПУ всех предметных заголовков, каждый из которых содержит все образующие поисковое предписание термины. Основная причина этого важнейшего недостатка АПУ состоит в следующем.

При индексировании документов присваиваемые им поисковые образы (предметные заголовки) могут состоять из любого числа терминов и иметь разную структуру. С другой стороны, информационный поиск обычно производится по сочетаниям не всех, а лишь некоторых терминов, использованных в поисковых образах документов. Поэтому предметные заголовки, в которых содержатся все образующие поисковое предписание термины, будут рассеяны в алфавитном ряду АПУ. Это объясняется тем, что в предметных заголовках между терминами, образующими поисковое предписание, могут стоять другие термины, не входящие в это поисковое предписание. Именно из-за этих терминов интересующие потребителя предметные заголовки попадают в разные места алфавитного ряда АПУ. Это означает, что при поиске документов по некоторому сочетанию терминов приходится использовать не один, а несколько предметных заголовков АПУ. Такие ситуации возникают даже тогда, когда применяются сравнительно широкие предметные заголовки.

Пусть дано следующее множество предметных заголовков АПУ:

| | |
|---|---|
| *ABF* | *ACH* |
| *ABH* | *ADE* |
| *ACE* | *ADG* |
| *ACF* | *ADH*, |

где $A$, $B$, $C$, $D$, $E$, $F$, $G$ и $H$ — отдельные термины. По этим заголовкам требуется найти все документы, предмет которых выражается сочетанием $AH$.

В приведенных выше предметных заголовках термины $A$ и $H$ входят в заголовки $ABH$, $ACH$ и $ADH$. Но в этих заголовках между терминами $A$ и $H$ стоят соответственно термины $B$, $C$ и $D$, из-за которых заголовки $ABH$, $ACH$ и $ADH$ попадают в разные места алфавитного ряда АПУ. Таким образом, в данном АПУ описания (или адреса хранения) искомых документов будут находиться под тремя разными заголовками — $ABH$, $ACH$ и $ADH$.

Очевидно, что АПУ будет эффективным лишь тогда, когда потребитель по любому сочетанию терминов сможет находить описания или адреса хранения всех интересующих его документов в одном месте указателя (т. е. под одним предмет-

ным заголовком). Для обеспечения этого каждый предметный заголовок в АПУ нужно представить в стольких вариантах, сколько перестановок могут дать образующие этот заголовок термины ($P_n = n!$, где $n$ — число терминов, образующих предметный заголовок).

Но если фиксировать последовательность расстановки терминов в предметном заголовке (естественным является алфавитный порядок), то тогда вместо всех возможных перестановок для каждого предметного заголовка достаточно составить все возможные сочетания из терминов, образующих этот заголовок и взятых по их любому числу.

Пусть $n$ — число терминов, образующих предметный заголовок. Тогда

$$\sum_{m=1}^{m=n} C_n^m = C_n^1 + C_n^2 + C_n^3 + \ldots + C_n^n = 2^n - 1.$$

Для сравнения полезно указать, что при $n=4$ суммарное число всех возможных сочетаний составляет 15, а число перестановок — 24.

В 1964 г. английский специалист по научной информации Дж. Шарп разработал вариант АПУ, для которого вполне достаточно значительно меньшего числа различных вариантов предметных заголовков. Этот указатель получил название SLIC (Selective Listing in Combination — Избирательное перечисление сочетаний) [45].

Общие принципы отбора предметных заголовков для такого АПУ могут быть раскрыты на следующем примере. Пусть дан документ $d_j$, основное смысловое содержание которого описано сочетанием терминов $A$, $B$, $C$ и $D$. Тогда из этих терминов можно построить следующие сочетания (заголовки), соблюдая в них алфавитный порядок расположения терминов:

*A, AB, ABC, ABCD, ABD, AC, ACD, AD, B, BC, BCD, BD, C, CD, D.*

Анализ этих сочетаний (заголовков) показывает, что некоторые из них представляют собой усеченные справа варианты более полных заголовков. Например, заголовки $A$, $AB$ и $ABC$ — суть усеченные варианты заголовка $ABCD$; заголовок $AC$ — усеченный вариант заголовка $ACD$ и т. д. Это означает, что для поиска документа $d_j$, например, по заголовкам $A$, $AB$ или $ABC$ можно пользоваться лишь неусеченным заголовком $ABCD$, который находится в том же месте алфавитного ряда, где бы мы искали заголовки $A$, $AB$ и $ABC$.

[45] J. P. Sharp. The SLIC Index. «American Documentation», 1966, v. 17, № 1, стр. 41—44. J. P. Sharp. Some fundamentals of information retrieval. London, Deutsch, 1965, стр. 81—87, 205—213.

Таким образом, из перечисленных выше 15 сочетаний (заголовков), составленных для документа $d_j$, можно без какого-либо ущерба для эффективности АПУ опустить заголовки

$$A,\ AB,\ ABC,\ AC,\ B,\ BC,\ C.$$

Оставшиеся восемь заголовков — $ABCD$, $ABD$, $ACD$, $AD$, $BCD$, $BD$, $CD$ и $D$ — обеспечивают выдачу документа $d_j$ независимо от того, какое сочетание из терминов $A$, $B$, $C$ и $D$, взятых по $m$ ($m$=1, 2, 3, 4), будет использовано для поиска этого документа. Общее число таких заголовков для документа, поисковый образ которого состоит из $n$ элементов, составляет $2^{(n-1)}$. Очевидно, что $2^{(n-1)} \ll (2^n-1)$.

На табл. 22 показано, насколько быстро растет число различных вариантов предметных заголовков для АПУ с увеличением $n$ ($n$ — число терминов в каждом предметном заголовке).[46]

Таблица 22

**Рост числа вариантов предметных заголовков в зависимости от числа терминов в каждом из них**

| $n$ | $P_n$ | $\sum_{m=1}^{m=n} C_n^m$ | $SLIC$ | $n$ | $P_n$ | $\sum_{m=1}^{m=n} C_n^m$ | $SLIC$ |
|---|---|---|---|---|---|---|---|
| 2 | 2 | 3 | 2 | 7 | 5040 | 127 | 64 |
| 3 | 6 | 7 | 4 | 8 | 40 320 | 255 | 128 |
| 4 | 24 | 15 | 8 | 9 | 362 880 | 511 | 256 |
| 5 | 120 | 31 | 16 | 10 | 3 628 800 | 1023 | 512 |
| 6 | 720 | 63 | 32 | | | | |

Итак, алфавитно-предметная классификация в ее каноническом виде, т. е. с выделением заголовков, подзаголовков и предметных записей, не позволяет производить информационный поиск непосредственно по сочетанию (логическому произведению) нескольких предметных заголовков (терминов). Вход в АПУ возможен только по одному заголовку (термину), после чего требуется последовательный просмотр некоторого множества подзаголовков и предметных записей. Поэтому для ответа на вопрос, насколько целесообразно применение алфавитно-предметной классификации в создаваемой ИПС, необходимо оценить среднее время, которое предположительно придется затрачивать при информационном поиске на просмотр каждого множества заголовков, начинающихся с данного слова, словосочетания или фразы. Далее это время необходимо сопоставить со средним числом информационных за-

[46] J. P. Sharp. Указ. соч., стр. 84.

просов, которые предположительно будут поступать в ИПС за день или неделю.

Обозначим через $M_i$ множество составных заголовков АПУ (т. е. цепочек вида «заголовок—подзаголовок—предметная запись»), начинающихся с $i$-того заголовка предметного словника. $M_i=(m_{i1}, m_{i2}, m_{i3}, \ldots, m_{ik})$, где $k$ — число элементов в $M_i$. Пусть $\tau$ — среднее время поиска любого элемента в $M_i$. Очевидно, что $\tau=f(k)$. Поэтому для оценки величины $\tau$ достаточно определить примерную величину $k$ при разных значениях таких параметров АПУ, как объем предметного словника, предполагаемое число документов в поисковом массиве и среднее число заголовков на документ.

С этой целью введем следующие дополнительные обозначения:

$V$ — общее число заголовков в предметном словнике (в данном случае имеются в виду цепочки «заголовок—подзаголовок»);

$C$ — предполагаемое число документов в поисковом массиве;

$p$ — среднее число заголовков (предметных записей), присваиваемых каждому документу при предметизации[47] $(1 \leqslant p \leqslant 10)$.

Тогда

$$k=\frac{Cp}{V}$$

(мы исходим из предположения, что все предметные заголовки употребляются равновероятно).

Величина $k$ зависит главным образом от численных значений $C$ и $V$, причем обычно $C \gg V$. Если $k$ достаточно велико (например, более $10^2$), тогда $\tau$ будет очень большим. В таких условиях алфавитно-предметная классификация в ее каноническом виде может оказаться неприемлемой как ИПЯ для документальной ИПС.

При возникновении такой ситуации необходимо изучить возможность значительного увеличения объема предметного словника[48]. Другой путь — приведение описаний документов (или адресов их хранения) под всеми возможными сочетаниями или даже перестановками терминов, из которых образованы заголовки, подзаголовки и предметные записи АПУ. В предельном случае $k=1$ и $\tau \to min$. Но для этого на каждый документ пришлось бы составлять $p(2^n-1)$, $p2^{n-1}$ или $pn!$ различных вариантов заголовков ($n$ — число попарно различных терминов в каждом заголовке).

---

[47] Н. М. Сагалович удачно назвал этот параметр «коэффициентом размножения».

[48] Методика оптимизации структуры АПУ детально разработана Н. М. Сагаловичем (см. его статьи: Уточнение терминологии, связанной с предметными указателями к реферативным изданиям. «Научно-техн. информация», 1966, № 11, стр. 21—24; Оптимизация структуры предметного указателя к реферативному журналу. «Научно-техн. информация», 1967, серия 2, № 5, стр. 3—4).

Очевидно, что данный путь не всегда приемлем, так как даже при сравнительно небольших значениях $V$, $p$ и $n$ это может потребовать весьма существенных расходов на составление предметных карточек (если не применяются соответствующие средства механизации) и привести к чрезмерному увеличению физического объема указателя или каталога.

Из всего сказанного следует, что если потребитель формулирует поисковые предписания в тех же терминах или в той же форме, в какой сформулированы предметные заголовки, то алфавитно-предметная система дает хорошие результаты. Но это требование трудно выполнимо, так как потребитель обычно не знает предметного словника и всех тонкостей его использования. Поэтому поиск документов по какому-либо предмету при помощи АПУ во многом напоминает поиск слова, например, в англо-русском словаре, когда неизвестно точное написание этого слова.

Опыт свидетельствует о том, что алфавитно-предметные системы более удобны для классификации не документов с их многоаспектным содержанием, а технических данных, различных изделий и т. д.

Анализ и обобщение изложенных выше сведений позволяет сформулировать основные достоинства и недостатки алфавитно-предметных классификаций как одного из типов ИПЯ. К числу важнейших достоинств таких классификаций относятся следующие:

1. Алфавитный порядок, в котором располагаются названия предметов (предметные заголовки), можно считать общеизвестным. Этот порядок широко применяется в энциклопедиях, словарях, телефонных справочниках и т. д. Поэтому для его использования не требуется никаких специальных знаний или вспомогательных средств. В этом заключается одно из основных преимуществ алфавитно-предметных классификаций перед различными ББК с их более или менее сложной структурой и нотацией.

2. Алфавитно-предметная классификация применима для любых физических форм ручных каталогов и указателей.

3. Для составления и использования алфавитно-предметного каталога или указателя не требуется никаких специальных средств и механизмов, хотя процедура алфавитно-предметного индексирования в принципе поддается механизации (например, по методу KWIC).

4. Существует много разных условий, в которых информационный поиск при помощи алфавитно-предметных каталогов и указателей может быть эффективным и экономичным без применения каких-либо средств механизации.

5. В алфавитно-предметные классификации значительно легче вводить новые термины, чем в ББК.

6. Алфавитно-предметная классификация в принципе может быть применена для индексирования неограниченного числа пред-

метов и документов, причем с увеличением числа предметов и документов не возникают трудности с их кодированием (шифровкой).

Суммируя основные достоинства алфавитно-предметных классификаций, Дж. Меткаф [49] писал:

«Знание алфавитного порядка означает больше, чем просто знание установленной последовательности фонетических знаков речи от *A* до *Z* и способность воспроизводить (recite) эту последовательность. Оно означает способность находить рубрики (entries) под названиями, расположенными в алфавитном порядке, причем для этого не нужны ни механические средства, ни последовательное сканирование или сплошной перебор этих названий.

Порядок следования букв и цифр в сочетаниях и перестановках, которые используются в качестве классификационных индексов, столь же широко известен. Такие сочетания и перестановки можно находить в списке даже быстрей, чем сочетания и перестановки букв, образующие названия. Но это не те буквенные сочетания и перестановки, написание и значение которых уже известно. Ручные или человеческие методы индексирования информации, которые основаны на использовании известных названий, располагаемых в известном порядке, во многих ситуациях будут иметь свои преимущества даже тогда, когда удастся разработать столь же эффективные машинные методы. Так, руки могут быть предпочтительными перед инструментами, ноги — перед колесами, тачки — перед моторными тележками для перевозки грузов, а книги, для чтения которых не нужно никаких других вспомогательных средств, кроме очков, — перед магнитофонами, аппаратами для чтения микроизображений и т. д.» [50].

Важнейшие недостатки алфавитно-предметных классификаций как ИПЯ состоят в следующем:

1. Как и ББК, алфавитно-предметные классификации не позволяют производить информационный поиск по любым сочетаниям предметов, обозначенных заголовками (т. е. по их логическим произведениям).

2. В каждой алфавитно-предметной классификации применяется какой-то один естественный язык, что затрудняет (или делает вообще невозможным) использование соответствующего каталога или указателя лицами, не владеющими этим языком. Преимуществом ББК перед алфавитно-предметными классификациями является то, что в них для обозначения классов обычно применяются символы, которые имеют международное распространение (например, арабские цифры).

---

[49] Дж. Меткаф — известный специалист по алфавитно-предметному индексированию. В настоящее время работает в университете Нового Южного Уэльса (Сидней, Австралия).

[50] J. Metcalfe. Alphabetical subject indication of information. (Rutgers series on systems for the intellectual organization of information, v.3). New Brunswick, N. J., The Rutgers University Press, 1965, стр. 15—16.

3. Включение в предметный словник синонимов и полусинонимов требует применения многочисленных перекрестных ссылок, что значительно усложняет структуру алфавитно-предметных каталогов и указателей.

4. Предметные словники алфавитно-предметных классификаций требуют значительно большего лексикографического контроля, чем таблицы ББК.

5. Словник алфавитно-предметной классификации не раскрывает потребителю основных правил построения этой классификации.

6. Составление алфавитно-предметных каталогов и указателей — весьма трудоемкий процесс, требующий больших затрат времени и усилий высококвалифицированных специалистов.

Как уже отмечалось выше, алфавитно-предметные каталоги и указатели предназначены главным образом для информационного поиска по предметам. Но они не пригодны тогда, когда преобладают информационные запросы тематического характера, особенно если при их формулировке исходят из какой-либо общепринятой классификации. Эту особенность нельзя отнести ни к достоинствам, ни к недостаткам алфавитно-предметной классификации, как нельзя считать достоинством орфографического словаря то, что он дает правильные написания слов данного языка, а его недостатком — что в нем не раскрывается значение этих слов.

Алфавитно-предметные классификации применяются в основном для индексирования книг и периодических изданий, включая реферативные журналы, а также для создания небольших каталогов и картотек специального назначения. Однако в последнем случае алфавитно-предметные классификации уже вытесняются координатным индексированием. Кроме того, алфавитно-предметные классификации применяются для составления каталогов к библиотечным фондам и указателей к систематическим каталогам.

Вопрос о применении алфавитно-предметных каталогов в библиотеках заслуживает особого внимания. Изучение этого вопроса показывает, что в библиотеках Советского Союза и других социалистических стран, а также в библиотеках стран Западной Европы алфавитно-предметные каталоги не получили широкого распространения. Исключение составляют лишь США, где сравнительно широкое применение алфавитно-предметных каталогов в библиотеках можно объяснить как чисто историческими причинами, так и прагматизмом, который столь характерен для американских библиографов и библиотековедов. Но и в библиотеках США систематические каталоги в последнее время начинают играть все более важную роль.

Основные причины, по которым библиотеки предпочитают иметь не алфавитно-предметные, а систематические каталоги, состоят в следующем:

а) по-видимому, большинство или значительная часть информационных запросов, с которыми приходится иметь дело в неспециальных библиотеках, носит тематический характер. Как было отмечено выше, алфавитно-предметные каталоги непригодны для информационного поиска по запросам такого типа;

б) разрешающая способность алфавитно-предметного каталога при информационном поиске по тематическим запросам может быть увеличена лишь путем введения в его структуру элементов классификации. Но это превратило бы данный каталог в нечто среднее между алфавитно-предметным и систематическим каталогом. Очевидно, что такой «гибридный» каталог будет значительно уступать систематическому каталогу при информационном поиске по тематическим запросам. С другой стороны, из-за значительного изменения и усложнения структуры алфавитно-предметного каталога снизится его эффективность и при информационном поиске по узкопредметным запросам;

в) если информационный поиск по немногочисленным запросам узкопредметного характера осуществляется при помощи систематического каталога, то для получения удовлетворительных результатов достаточно дополнить такой каталог алфавитно-предметным указателем. Составление алфавитно-предметного указателя к систематическому каталогу требует значительно меньших затрат времени и средств, чем составление отдельного алфавитно-предметного каталога. Таким образом, в условиях библиотеки значительно более эффективным оказывается систематический каталог с алфавитно-предметным указателем, чем модифицированный вариант алфавитно-предметного каталога;

г) большинство библиотек из-за нехватки средств не может позволить себе одновременное ведение обоих каталогов — систематического и алфавитно-предметного — на весь фонд имеющейся в библиотеке литературы. «Может быть, не в одинаковой степени, но наличие предметного каталога, наряду с систематическим, желательно во всех специальных библиотеках, — пишет З. Н. Амбарцумян. — Однако следует считаться не только с желаниями, но и с возможностями» [51].

Данный раздел целесообразно завершить рассмотрением достоинств и недостатков алфавитно-предметных указателей «цепного» типа. Ранее несколько раз подчеркивалось, что одна из основных трудностей, с которыми приходится сталкиваться при построении ИПС, состоит в невозможности предвидеть все варианты, в которых потребители будут формулировать свои информационные запросы. Например, поиск документов о зимнем опрыскивании пшеницы пестицидами может производиться по терминам «Зима», «Опрыскивание», «Пшеница» и «Пестициды». Поэтому очень

[51] З. Н. Амбарцумян. Некоторые спорные вопросы предметного каталога. «Сов. библиография», 1965, № 5, стр. 42.

важно, чтобы в алфавитно-предметном указателе к ББК (или к систематическому каталогу) в качестве ведущих выступали все термины, которые могут быть использованы при информационном поиске, т. е. чтобы указатель обладал определенной избыточностью, обеспечивающей нахождение любого класса независимо от того, по какому термину производится поиск этого класса.

Один из возможных методов для достижения этой цели состоит в представлении названия каждого класса в виде всех возможных перестановок слов, составляющих это название. Например, для приведенного выше названия («Зимнее опрыскивание пшеницы пестицидами») мы будем иметь следующие перестановки [52]:

Пестициды : пшеница : опрыскивание : зима
633.1-298.3
Пшеница : пестициды : опрыскивание : зима
633.1-298.3 и т. д.

Но данный метод не может быть использован без соответствующих ограничений, так как это привело бы к быстрому росту объема указателя. Действительно, для данного случая пришлось бы построить $4!=24$ вариантам названия этого предмета.

Наиболее удобный метод порождения вариантов названий классов для алфавитно-предметного указателя к ББК (или к систематическому каталогу) — это метод «цепного» индексирования, разработанный Ш. Р. Ранганатаном в начале 30-х годов. Процедура цепного индексирования состоит в следующем.

При отнесении документа к какому-либо классу ББК, его предмет выражается в виде цепочки названий подклассов. Эта цепочка начинается с названия главного класса и кончается названием наиболее узкого подкласса. Каждый такой подкласс по возможности обозначается одним термином. Обычно этот термин берется из таблиц ББК, хотя возможны определенные исключения из этого правила.

Допустим, что в документе говорится о машинах для уборки зерновых культур. Тогда предмет данного документа выражается следующей цепочкой названий классов и подклассов УДК:

| | |
|---|---|
| 6 | Техника |
| 63 | Сельское хозяйство |
| 633 | Полеводство |
| 633.1 | Хлебные злаки. Зерновые |
| 633.1-13 | Орудия. Машины |
| 633.1-135 | Уборка урожая |

Теперь путем систематического изучения названия каждого подкласса, начиная с последнего (633.1-135), можно построить названия всех вариантов этого класса для алфавитно-предмет-

[52] Guide to the Universal Decimal Classification (UDC). B. S. 1000C: 1963 (F. I. D. N. 345). London, British Standards Institution, 1963, стр. 34—41.

ного указателя. При этом каждый раз необходимо задавать вопрос, может ли быть использован данный термин в качестве исходного для обращения в указатель. Если на этот вопрос дается утвердительный ответ, то строится вариант заголовка, начинающийся с этого термина. Например, для последнего подкласса мы имеем вариант:

Уборка урожая : Машины : Зерновые 633.1-135

Из приведенного выше примера видно, что ведущий термин «Уборка урожая» снабжается соответствующими пояснительными словами. Общее правило применения таких пояснительных слов состоит в том, что к ведущему термину добавляются лишь такие пояснительные слова, которые необходимы для устранения неоднозначности этого ведущего термина. Опускаются такие пояснительные слова, которые в той или иной степени уже выражены в других пояснительных словах. В данном случае опущены слова «Полеводство», «Сельское хозяйство» и «Техника».

Второе правило требует, чтобы к ведущему термину добавлялись лишь такие пояснительные слова, которые в иерархической цепочке стоят выше данного ведущего термина.

Таким образом, для предмета «Машины для уборки зерновых культур» мы имеем следующее множество заголовков:

| | |
|---|---|
| Уборка урожая : Машины : Зерновые | 633.1-135 |
| Машины : Зерновые : Сельское хозяйство | 633.1-13 |
| Орудия : Зерновые : Сельское хозяйство | 633.1-13 |
| Зерновые : Сельское хозяйство | 633.1 |
| Хлебные злаки : Сельское хозяйство | 633.1 |
| Полеводство | 633 |
| Сельское хозяйство | 63 |

Механизм построения «цепного» указателя можно еще более наглядно показать на примере фасетной классификации [53]. Допустим, что основное смысловое содержание документа сформулировано следующим образом: «Расчет ламинарного пограничного слоя на треугольном крыле». Тогда при использовании фасетной классификации этот документ будет отнесен к классу, имеющему следующее название:

Крылья : Треугольные : Ламинарный поток : Пограничный слой : Расчет
*CdIjNcdNfkYa*

Для составления алфавитно-предметного указателя методом цепного индексирования порядок терминов в этом названии должен быть изменен на обратный. Новое название с обратным порядком слов становится заголовком алфавитно-предметного

[53] B. C. Vickery. Faceted classification. (Rutgers series on systems for the intellectual organization of information. V. 5). New Brunswick, N. J., The Rutgers University Press, 1966, стр. 77—78.

указателя к ББК (или к систематическому каталогу). Этот заголовок будет иметь следующий вид:

Расчет : Пограничный слой : Ламинарный поток :
: Треугольные : Крылья

Затем первый (ведущий) термин этой цепочки отбрасывается, и ведущим становится второй термин (в шифре класса отбрасывается последний символ —*Ya*). Это усеченное сочетание терминов также становится заголовком цепного указателя. Такая процедура повторяется до тех пор, пока ведущим не станет последний термин сочетания. В результате этого мы получим следующее множество рубрик:

| | |
|---|---|
| Расчет : Пограничный слой : Ламинарный поток : : Треугольные : Крылья | *CdIjNcdNjkYa* |
| Пограничный слой : Ламинарный поток : Треугольные : : Крылья | *CdIjNcdNjk* |
| Ламинарный поток : Треугольные : Крылья | *CdIjNcd* |
| Треугольные : Крылья | *CdIj* |
| Крылья | *Cd* |

Цепной метод построения алфавитно-предметного указателя к ББК (или к систематическому каталогу) имеет следующие основные достоинства:

1) цепной метод представляет собой в значительной степени механическую процедуру, при последовательном осуществлении которой опасность пропуска полезного варианта названия рубрики сведена до минимума;

2) цепной метод является весьма экономичным, так как благодаря строгому контролю за последовательностью расположения слов в вариантах названия класса очень сильно сокращается число допустимых перестановок этих слов, причем это не уменьшает эффективности указателя.

Однако «цепные» указатели к ББК или систематическим каталогам имеют в то же время ряд существенных недостатков. Например, не трудно понять, что чем шире значение термина, используемого для обращения в такой указатель, тем больший раздел ББК или систематического каталога нужно будет просмотреть для отыскания интересующих документов. Кроме того, поиск приходится производить во многих разделах ББК (или систематического каталога), так как термин, выбранный для обращения в указатель, может употребляться в самых разных сочетаниях с другими терминами, причем некоторые из этих сочетаний могут в большей или меньшей мере соответствовать информационному запросу.

Кроме того, «цепной» указатель, который до последнего времени считался неотъемлемым атрибутом фасетных классификаций, значительно снижает эффективность таких классификаций. Это может быть показано на следующем примере.

Пусть дан документ, предмет которого в терминах некоторой фасетной классификации описан выражением *ABCDEF*. Тогда это выражение в «цепном» указателе будет представлено следующим множеством заголовков:

*FEDCBA*
*EDCBA*
*DCBA*
*CBA*
*BA*
*A*

Как было указано выше, основное назначение «цепного» указателя к систематическому каталогу состоит в том, чтобы обеспечить быстрое отыскание в этом каталоге рубрик, заголовки которых содержат любой термин в любом заданном контексте, т. е. в сочетании с любым другим термином или терминами.

Стратегия поиска сочетания из двух терминов (например, сочетания терминов *C* и *E*) в систематическом каталоге состоит в следующем.

1. По некоторому термину поискового предписания, выбранному в качестве исходного, мы отыскиваем в «цепном» указателе к систематическому каталогу все заголовки, начинающиеся с этого термина. Затем, последовательно прочитывая эти заголовки, мы выявляем такие из них, которые одновременно содержат и другой заданный термин. Но если в данном случае начать поиск с термина *C*, то *в заголовке «цепного» указателя, начинающегося с C, мы не обнаружим термина E*. Это объясняется тем, что в выражении *ABCDEF* термин *C* стоит перед *E*. Поэтому при составлении «цепного» указателя этот термин будет отброшен до того, как *C* станет ведущим термином.

2. При возникновении такой ситуации можно воспользоваться следующей альтернативой:

а) повторить поиск сочетания *CE*, используя в качестве исходного термин *E*;

б) просмотреть все рубрики систематического каталога, на которые сделаны отсылки от заголовка, начинающегося с термина *C*. Если отсутствие в «цепном» указателе заголовков, начинающихся с *C* и одновременно содержащих *E*, объясняется причиной, которая была указана выше в п. 1, то в систематическом каталоге будут найдены искомые рубрики.

Сравнительно низкая эффективность фасетных классификаций при поиске информации подтверждается результатами Крэнфилдского проекта. Это объясняется главным образом фиксацией последовательности расположения терминов в фасетной формуле. Поэтому высказываются мнения, что необходимо отказаться от фиксированной последовательности терминов в фасетной формуле и применять все возможные перестановки терминов.

и морфологические показатели, выполняют грамматические функции актуализации полнозначимых слов, их сплочения в единое предложение, модальной характеристики, высказываний и выражения межфразовых связей»[2]. К полнозначимым словам относится подавляющее большинство существительных, прилагательных, глаголов, наречий, числительных и местоимений, а к неполнозначным — предлоги, союзы, связки, частицы и многие другие слова.

Полнозначимая лексика состоит из назывных слов и местоимений, которые не обозначают предмет, а выделяют его путем указаний (дейксиса). К категории назывных слов относятся также имена собственные. Некоторые языковеды считают имена собственные идеальными названиями предметов. С. Д. Кацнельсон указывает на следующие признаки назывных слов, которые противопоставляют их деиктическим словам:

1) надситуативность, т. е. выделение определенного объекта независимо от того, присутствует он в данный момент в чувственной ситуации или нет;

2) неэгоцентричность (выбор слова не зависит от говорящего лица, его местоположения и времени высказывания);

3) объективность (выделение предмета достигается теперь в опоре на некие релевантные признаки самого предмета);

4) семантическая устойчивость и контрастность (в каждом акте употребления они сохраняют некое инвариантное ядро, определенный минимум признаков, необходимых для распознания предмета)[3].

Далее С. Д. Кацнельсон пишет: «Объективный характер назывных слов роднит их с понятиями. Лексические значения, как и понятия, это своего рода умственные «концентраты», сгустки человеческих знаний об определенных фрагментах и сторонах окружающей нас действительности. Значения слов, как и понятия, покоятся на специфической форме отражения действительности — обобщении и абстракции»[4].

Необходимо отметить, что, кроме назывных слов, в функции ключевых слов могут выступать также соответствующие численные характеристики, хронологические данные, диапазоны температур, давлений и т. д. Ниже приведен пример координатного индексирования реферата 6Г180, помещенного в РЖ «Физика» за 1961 г.[5]

---

[2] С. Д. Кацнельсон. Содержание слова, значение и обозначение. М.—Л., «Наука», 1965, стр. 3.

[3] Там же, стр. 6.

[4] Там же, стр. 9.

[5] Пример взят из американского реферативного журнала «Abstracts of Photographic Science and Engineering Literature», 1962, v. 1, № 3—4, стр. 68—69.

5Г180. **Проявляемость фотографического скрытого изображения, созданного действием γ-излучения.** Томода И., Оцу Е. «Токё когё сикэнсё хококу, Repts Govt Chem. Industr. Res. Inst., Tokyo», 1960, 55, № 7, 235—240, 26 (японск.; рез. англ.). — Исследована кинетика проявления почернений, образованных на кинопозитивной эмульсии действием γ-излучения $Co^{60}$ и, для сравнения, света; в последнем случае экспонирование велось при нескольких весьма различных выдержках ($10^2$; $5 \cdot 10^{-2}$ и $10^{-7}$ сек.) и различных освещенностях. Проявление в метолгидрохиноновом проявителе D-72 характеризовалось после действия γ-излучения относительно малой скоростью, даже если окончательное почернение велико. Такая особенность наблюдалась и для очень кратковременной засветки и соответствует обычному эффекту Кабанна—Гофмана, т. е. очень высокой дисперсности скрытого изображения. При проявлении глубинным проявителем Стивенса различия кинетики для действия γ-излучения и засветки любой длительности исчезали, что, согласно авторам, указывало на значительное влияние скорости растворения AgHal на кинетику проявления. Это непосредственно подтвердилось опытами по изменению конц-ии растворителей AgHal в проявителе. Измерение температурных коэфф. проявления для проявителя D-72 в интервале 5—25° С показало, что энергия активации проявления (определяемая по ур-нию Аррениуса с использованием времени проявления для достижения заданной плотности в качестве критерия) почти одинакова для действия γ- и светового излучения (даже при больших выдержках) и зависит гораздо сильнее от состава проявителя, чем от дисперсности скрытого изображения. *А. Картужанский*

галогениды серебра

гамма-излучение

кинетика

плотность

проявление

скрытое изображение

экспонирование

эффект Кабанна—Гофмана

Приведенный справа от реферата алфавитный список ключевых слов образует поисковый образ этого реферата. Практический опыт показывает, что для координатного индексирования одного документа обычно бывает достаточно 8—12 ключевых слов.

Итак, координатное индексирование — это способ выражения основного смыслового содержания (предмета) документа или информационного запроса в виде определенной совокупности ключевых слов. Координатное индексирование называют также *методом координации понятий*, *коррелятивным индексированием*, *манипулятивным индексированием*, *унитерм-индексированием*, *ассоциативным индексированием*, *комбинаторным индексированием*, *зато-кодированием и т. д.*

При «чистом» координатном индексировании ключевые слова в поисковых образах никак не связаны одно с другим и функционируют самостоятельно. В отличие от поисковых образов документов поисковые предписания формулируются в виде логических сумм, произведений или дополнений классов, обозначаемых соответствующими ключевыми словами. Для отыскания

документов, отвечающих на какой-либо информационный запрос, необходимо выполнить определенные логические операции над классами, которые обозначены ключевыми словами поисковых образов документов. В простейшем случае, когда поисковое предписание сформулировано в виде логического произведения (конъюнкции) некоторого множества ключевых слов, документ считается отвечающим на информационный запрос и подлежит выдаче, если в поисковом образе этого документа одновременно содержатся все ключевые слова поискового предписания. В определенном смысле это эквивалентно выполнению операции логического умножения классов, обозначенных теми ключевыми словами поискового образа, которые совпадают с ключевыми словами, образующими поисковое предписание. «Координатное индексирование, — пишут М. Таубе и Г. Вустер, — это такой метод анализа элементов информации, при котором их поиск осуществляется путем логических операций умножения, сложения и дополнения над кодами, записанными в запоминающем устройстве (store)» [6].

Очевидно, что для обеспечения единообразия в координатном индексировании документов и информационных запросов необходимо лишь условиться об однозначном понимании используемых для этого ключевых слов. Это значительно более простая задача, чем попытки договориться об одинаковом понимании основного смыслового содержания (предмета) того или иного документа. Точно так же значительно легче прийти к соглашению относительно того, какие ключевые слова необходимы и достаточны для выражения основного смыслового содержания документа, чем о предметных заголовках (одном или нескольких), которые адекватно индексируют этот документ.

Внешне система координатного индексирования как будто очень похожа на алфавитно-предметную классификацию. Действительно, в обоих случаях мы имеем дело с классами, которые обозначаются соответствующими словами и словосочетаниями естественного языка, т. е. ключевыми словами и предметными заголовками. Например, ключевое слово «электрический» обозначает целый класс предметов и понятий, названия которых содержат это слово. При формулировании предметных заголовков также порождаются соответствующие классы предметов и понятий. Но этим и исчерпывается сходство координатного индексирования с алфавитно-предметной классификацией.

Выше было указано, что для отыскания документов, поисковые образы которых составлены из определенных совокупностей ключевых слов, необходимо производить соответствующие логические операции над классами, обозначенными этими ключевыми

[6] M. Taube and H. Wooster (Eds). Information storage and retrieval. Theory, systems and devices. New York, Columbia University Press, 1958, стр. 8.

словами. Ничего подобного не требуется при информационном поиске с помощью алфавитно-предметных каталогов (указателей).

Для алфавитно-предметного индексирования одного документа также может быть применено несколько предметных заголовков, каждый из которых будет исчерпывающе характеризовать один из предметов всего документа (текста) или его определенной части. Но при информационном поиске каждый предметный заголовок выступает самостоятельно, без какой-либо связи с другими заголовками. Адресные шифры всех искомых документов должны находиться под соответствующим предметным заголовком. Поэтому для отыскания нужных документов не требуется одновременного использования нескольких предметных заголовков и тем более их логического умножения, сложения или дополнения. Аналогичная ситуация в большей или меньшей степени наблюдается и при использовании как иерархических, так и фасетных классификаций.

Преимущество координатного индексирования перед алфавитно-предметной классификацией можно показать на следующем примере. Предположим, что для индексирования какого-либо документа было использовано 5 ключевых слов. Если бы мы захотели обеспечить возможность поиска этого документа по любому сочетанию из этих 5 ключевых слов в алфавитно-предметном каталоге (указателе), то лишь для индексирования этого документа нам нужно было бы иметь в данном каталоге (указателе) $P_5 = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120$ предметных заголовков.

Одним из вариантов координатного индексирования является так называемое «глубокое индексирование» (deep indexing), разработанное Дж. Костелло (США) в 1962 г.[7] Оно применяется в системах, предназначенных для узких групп специалистов, информационные потребности которых хорошо изучены и четко сформулированы, или для использования в каких-либо специфических целях. Основное различие между обычным и глубоким координатным индексированием заключается в следующем. Обычное координатное индексирование, как правило, основано на более или менее поверхностном изучении текста документов. Предметизатор не определяет целевого назначения индексируемого им документа и не оценивает содержащуюся в нем информацию в отношении текущих и предвидимых потребностей обслуживаемой группы специалистов и в сравнении с другими документами поискового массива.

При глубоком координатном индексировании предметизатор производит тщательный анализ информации, содержащейся в

[7] J. C. Costello, Jr. An introduction to deep (coordinate) indexes. «Journal of Chemical Documentation», 1963, v. 3, № 3, стр. 164—168; J. C. Costello, Jr. Indexing in depth: practical parameters. В кн.: P. Howerton (Ed.). Information handling: first principles. Washington, Spartan Books, 1963, стр. 55—87.

индексируемом документе, определяет ее относительную ценность для обслуживаемой организации или группы потребителей и в зависимости от этого более или менее подробно (глубоко) индексирует данный документ. Глубина индексирования определяется также задачами организации или группы, для которой создана поисковая система, и такими факторами, как опыт всей группы потребителей или отдельных ее членов, их образование и т. д. Дж. Костелло считает, что глубокое координатное индексирование в большей степени основано на целевом назначении информации, содержащейся в документах (use-oriented), а также на ценности этой информации (value-oriented) и в меньшей степени — на словах, используемых в тексте документа (word-oriented). Очевидно, что глубокое индексирование может производить лишь специалист в той отрасли науки или техники, к которой относятся индексируемые документы.

В связи с этим необходимо отметить, что глубина координатного индексирования какого-либо документа не измеряется числом ключевых слов, включенных в поисковый образ этого документа. Это можно показать на следующем примере. Предположим, что поисковый образ документа состоит из ключевых слов $A$, $B$, $C$ и $D$. Число ключевых слов в этом поисковом образе можно увеличить вдвое, если включить в него ключевые слова $A'$, $B'$, $C'$ и $D'$, которые в иерархическом отношении стоят непосредственно выше ключевых слов $A$, $B$, $C$ и $D$. Но такое удвоение числа ключевых слов в поисковом образе документа ни на йоту не увеличивает глубины координатного индексирования этого документа. Глубина координатного индексирования определяется главным образом степенью полноты отражения в поисковом образе документа всех наиболее важных специфических понятий, рассматриваемых в этом документе и выражаемых в его поисковом образе при помощи наиболее специфических (узких) ключевых слов.

Ниже дается формальное описание основных принципов «чистого» координатного индексирования в элементарных терминах теории множеств. Но для этого сначала определим важнейшие понятия и термины этой теории. *Множество* — это просто совокупность каких-либо элементов. Буквы алфавита, статьи в журналах, книги на полках, числа 1, 2, 3, 4, 5 — все это примеры множеств. Множество может содержать только один элемент или совсем не иметь элементов (*пустое множество*). Оно также может быть бесконечным (например, бесконечно число точек на окружности). Множество обычно обозначают написанием его элементов внутри фигурных скобок, например {1, 2, 3, 4, 5}.

Запись высказываний о множествах и их элементах обычно производится при помощи следующих общепринятых сокращений. Высказывание «5 является элементом множества $B$» записывается $5 \in B$; высказывание «7 не принадлежит множеству $B$» записывается $7 \notin B$ или $7 \bar{\in} B$. Высказывание «$A$ есть подмножество $B$»

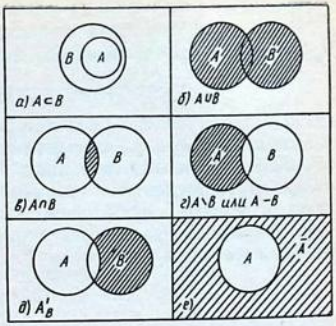



Рис. 43. Графическое изображение соотношений между множествами при помощи фигур Венна

а) *включение*: $A$ является подмножеством $B$; б) *объединение* (или *сумма*) $A$ и $B$ (заштриховано); в) *пересечение* (или *произведение*) $A$ и $B$ (заштриховано); г) *разность* $A$ и $B$ (заштрихована); д) *дополнение* $A$ и $B$ (заштриховано); е) *дополнение* $A$ в универсальное множество (заштриховано)

(т. е. такое, все элементы которого входят в множество $B$) записывается $A \subseteq B$ или $B \supseteq A$. Если $A$ есть подмножество $B$, причем $A \neq B$, то пишут $A \subset B$ или $B \supset A$. Пустое множество обозначается символом $\varnothing$ или 0. Множество, которое содержит все объекты, рассматриваемые в данной области, называется *универсальным множеством* и обозначается буквой $U$ или $I$.

Основными действиями над множествами являются *объединение* (или *сложение*), *пересечение* (или *умножение*), *вычитание* и *дополнение*. Объединением или суммой двух множеств $A$ и $B$ (записывается $A \cup B$) называется новое множество, состоящее из всех элементов $A$ и всех элементов $B$. Таким образом, объединением (суммой) множеств $\{1, 2, 3\}$ и $\{2, 3, 4, 5\}$ будет множество $\{1, 2, 3, 4, 5\}$. Пересечением или произведением двух множеств $A$ и $B$ (записывается $A \cap B$) называется множество всех элементов, которые являются общими для этих обоих множеств. Пересечением (произведением) множеств $\{1, 2, 3\}$ и $\{2, 3, 4, 5\}$ является множество $\{2, 3\}$. Разностью двух множеств $A$ и $B$ (записывается $A \setminus B$ или $A - B$) называется такое множество, в которое входят все элементы множества $A$, не принадлежащие множеству $B$. Разностью множеств $\{1, 2, 3\}$ и $\{3, 4, 5\}$ является множество $\{1, 2\}$. Наконец, дополнением множества $A$ в $B$ (записывается $A/B$) является множество всех элементов $B$, которые не являются элементами множества $A$. Дополнением множества $\{1, 2, 3\}$ в множество $\{2, 3, 4, 5\}$ будет множество $\{4, 5\}$. Дополнение множества $A$ в универсальное множество записывается как $\bar{A}$ или $A'$. Графическое изображение основных соотношений между множествами дано на рис. 43 [8].

[8] П. С. Александров. Введение в теорию множеств и функций. М.—Л., Гостехиздат, 1948; В. Серпинский. О теории множеств. М., «Просвещение», 1966.

Пусть дано некоторое множество документов $A$, в котором производится информационный поиск,

$$A = \{a_1, a_2, a_3, \ldots, a_n\},$$

где $a_1, a_2, a_3, \ldots, a_n$ — отдельные документы (или их номера). Множество $A$ будем рассматривать как универсальное множество, заданное точным перечислением всех его элементов. Критерием отбора из $A$ документов с требуемой информацией является вхождение в тексты, выражающие основное смысловое содержание этих документов, некоторого множества ключевых слов, взятых из универсального множества $T$.

$$T = \{t_1, t_2, t_3, \ldots, t_m\},$$

где $t_1, t_2, t_3, \ldots, t_m$ — отдельные ключевые слова. Множество $T$ также может быть задано точным перечислением всех элементов.

При соотнесении $A$ и $T$ любое ключевое слово $t_i$, взятое из множества $T$, вычленяет (обозначает) в $A$ подмножество документов $A(t_i)$, в котором текст, выражающий основное смысловое содержание каждого документа, содержит это ключевое слово $t_i$ (или обладает свойством $t_i$). Множество $A(t_i)$ может быть одноэлементным или даже пустым (нулевым). $A(t_i) = \varnothing$, если все элементы $T$ заданы точным перечислением до ввода документов в ИПС. Тогда в начальный период функционирования документальной ИПС в его запоминающем устройстве (ЗУ) может еще не быть ни одного документа, которому бы соответствовало данное ключевое слово $t_i$.

Очевидно, что если всякое ключевое слово $t_i \in A$ вычленяет свое подмножество документов $A(t_i)$, то набор из $\alpha$ таких ключевых слов также образует свое подмножество документов ($1 \leqslant \alpha \leqslant m$). Это множество есть пересечение множеств, обозначаемых соответственно ключевыми словами $t_{i1}, t_{i2}, t_{i3}, \ldots, t_{i\alpha}$ [9].

Если ключевые слова $t_{i1}, t_{i2}, t_{i3}, \ldots, t_{i\alpha}$ образуют соответственно множества $A(t_{i1}), A(t_{i2}), A(t_{i3}), \ldots, A(t_{i\alpha})$, то набор $\{t_{i1}, t_{i2}, t_{i3}, \ldots, t_{i\alpha}\}$ образует из элементов $A$ некоторое новое множество $A(T_i)$, которое есть пересечение множеств $A(t_{i1})$, $A(t_{i2})$, $A(t_{i3})$, $\ldots$, $A(t_{i\alpha})$:

$$A(T_i) = \bigcap_{\alpha} A(t_{i\alpha}),$$

где $T_i = \{t_{i1}, t_{i2}, t_{i3}, \ldots, t_{i\alpha}\}$ и $A(T_i) \subset A$.

[9] Если принадлежность любого ключевого слова $t_i$ к множеству $T$ интерпретировать как суждение вида $t_i \in T$, то тогда набор из $\alpha$ таких суждений можно представить в виде конъюнкции (логического произведения) $t_{i1} \in T \& t_{i2} \in T \& t_{i3} \in T \& \ldots \& t_{i\alpha} \in T$; эта конъюнкция истинна в том, и только в том случае, когда истинны все составляющие ее простые высказывания, т. е. когда все ключевые слова $t_i$ — суть элементы множества $T$.

Пусть $\alpha = 3$. Тогда графически операцию пересечения $A(t_{i1})$, $A(t_{i2})$ и $A(t_{i3})$, которая производится при информационном поиске, можно представить в виде, показанном на рис. 44.

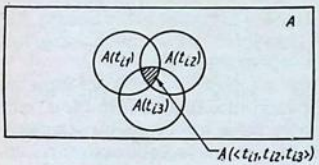



Рис. 44. Графическое изображение операции пересечения множеств $A(t_{i1})$, $A(t_{i2})$ и $A(t_{i3})$

Из сказанного выше следует, что при соотнесении множеств $A$ и $T$ каждому документу $a_i \in A$ ставится в однозначное соответствие определенный набор ключевых слов из $T$, которые выражают основное смысловое содержание данного документа $a_i$. Иными словами, имеет место отображение элементов множества $A$ в множество $T$.

Пусть $P$ — отображение множества $A$ в множество $T$. Это означает, что каждому $a_i \in A$ поставлено в однозначное соответствие подмножество ключевых слов $\{t_i\} = T_i$, $T_i \subset T$, которые входят (в явном или скрытом виде) в текст, выражающий основное смысловое содержание $a_i$.

$$a_i \to \{t_{i1},\ t_{i2},\ t_{i3},\ \ldots,\ t_{i\alpha}\} = T_i.$$

Очевидно, что в множестве $A$ может содержаться не один, а несколько документов, имеющих одинаковое отображение в $T$. Поэтому более уместно говорить об отображении в $T$ не отдельных документов из $A$, а соответствующих подмножеств таких документов.

Множество ключевых слов $PA(T_i) = \{t_{i1},\ t_{i2},\ t_{i3}, \ldots, t_{i\alpha}\} = T_i$, поставленное в однозначное соответствие каждому документу $a_i \in A(T_i)$, $A(T_i) \subset A$, выражающее (с приемлемой степенью точности и полноты) основное смысловое содержание этого документа и предназначенное для его последующего отыскания в $A$, может быть названо *поисковым образом* этого документа (или множества $A(T_i)$ таких документов). Операцию такого отображения элементов множества $A$ в множество $T$, т. е. описание основного смыслового содержания каждого документа списком соответствующих ключевых слов, называют *координатным индексированием*. При этом необходимо указать, что поисковый образ $PA(T_i)$ считается правильно построенным тогда, и только тогда, когда соблюдены следующие условия:

1) $PA(T_i)$ не содержит ни одного повторяющегося ключевого слова $t_i$;

2) $PA(T_i)$ не содержит ни одного ключевого слова $t_i$, которое не входит одновременно в каждый текст, выражающий основное смысловое содержание каждого документа множества $A$ $(T_i)$;

3) $PA(T_i)$ не содержит ни одного ключевого слова $t'$, которое не является элементом $T$.

Если множество $T$ задано точным перечислением всех его элементов (т. е. предварительно составлен список всех возможных ключевых слов), то в простейшем случае координатное индексирование состоит в последовательном сопоставлении слов текста, выражающего основное смысловое содержание индексируемого документа (или всего текста документа), с этим списком и выписывании из него (без повторений) тех ключевых слов, которые встречаются в данном тексте. Полученный список ключевых слов $T_i$, поставленный в соответствие этому документу, и будет его поисковым образом.

Ключевые слова могут быть не заданы предварительно составленным списком. Тогда ключевыми словами считаются такие, которые наиболее часто используются в индексируемом тексте и не входят в некоторый предварительно составленный список неключевых слов.

Известно, что всякая задача, для решения которой имеется алгоритм [10], в принципе может быть передана для решения машине. Поскольку любой из описанных выше процессов индексирования текстов, выражающих основное смысловое содержание документов, является алгоритмическим, т. е. может быть расчленен на конечную и совершенно однозначную последовательность элементарных дискретных операций, приводящую к преобразованию различных текстов в соответствующие им поисковые образы (множества ключевых слов), индексирование документов по описанной выше схеме может производиться машиной. Однако такое индексирование далеко не всегда будет адекватным.

Процедура информационного поиска в массиве документов, поисковые образы которых образованы путем «чистого» координатного индексирования этих документов, может быть формально описана следующим образом. Пусть $P_q$ — поисковое предписание, выражающее смысловое содержание информационного запроса $Q$.

$$P_q = (t_{i1}, t_{i2}, t_{i3}, \ldots, t_{is}), \quad t_i \in T.$$

Документ $a_i$ считается отвечающим на информационный запрос $Q$ и подлежит выдаче, если между поисковым образом $PA$ $(T_i)$

---

[10] Алгоритм описательно может быть определен как точное изложение содержания и последовательности некоторой детерминированной системы операций, переводящих варьируемые исходные данные в искомый результат.

этого документа и поисковым предписанием $P_q$ имеет место отношение $P_q \subseteq PA\ (T_i)$.

Вернемся теперь к содержательному рассмотрению «чистого» координатного индексирования. Пусть даны докуметы 02, 05 и 08, предметами которых являются:

02 — хлористый натрий, бромистый калий и иодистый калий.
05 — столкновение пи-мезона с протоном.
08 — передача электроэнергии из Англии в Шотландию.

При чистом координатном индексировании эти документы могут иметь следующие поисковые образы:

02 — *хлориды, натрий, бромиды, калий, иодиды.*
05 — *столкновения, пи-мезоны, протоны.*
08 — *передача, электроэнергия, Англия, Шотландия.*

Ниже на основании этих (и некоторых других) данных построены примеры, показывающие основные случаи недостаточности «чистого» координатного индексирования для обеспечения высокой точности и полноты информационного поиска. Эти примеры приводятся в целях иллюстрации, и поэтому им намеренно придан тривиальный характер.

1. *Ложная координация.* Предмет информационного запроса — «Иодистый натрий». Поисковое предписание сформулировано так: *иодиды, натрий.* Документальная ИПС выдает документ 02 (хлористый натрий, бромистый калий и иодистый калий). Причина выдачи нерелевантного документа — недостаточность лишь координативной связи между ключевыми словами в поисковом образе документа 02.

2. *Неполная координация.* Предмет информационного запроса — «Столкновение протона с протоном». Поисковое предписание сформулировано так: *столкновения, протоны.* Документальная ИПС выдает документ 05 (столкновение пи-мезона с протоном). Причина выдачи нерелевантного документа заключается в том, что для формулировки поискового предписания были использованы ключевые слова, достаточные для выражения предмета информационного запроса, но недостаточные для выражения предмета документа 05, хотя ИПС и выдала этот документ. Для предотвращения выдачи документа 05 необходимо сформулировать поисковое предписание следующим образом: *столкновения, протоны*, НЕ *мезоны.* Но для этого потребитель должен заранее знать о необходимости таких логических отрицаний.

3. *Синонимия ключевых слов.* Предмет информационного запроса — «Поваренная соль». Формулировка поискового предписания — *поваренная соль.* Документальная ИПС не выдает ни одного документа, хотя документ 02 отвечает на информационный запрос. Причина невыдачи релевантного документа заключается в том, что вместо ключевого слова *поваренная соль* в поисковом

образе документа 02 был использован его синоним «хлористый натрий», выраженный ключевыми словами *хлориды, натрий*. Для предотвращения невыдачи документа 02 нужно было устранить синонимию ключевых слов, применяемых для координатного индексирования документов и информационных запросов. Это означает, что такие ключевые слова (т. е. лексика применяемого ИПЯ) должны быть заданы в виде нормализованного списка.

4. *Полисемия ключевых слов.* Предмет информационного запроса — «Реакторы». Если в поисковом предписании не будет уточнено, какие реакторы интересуют в данном случае потребителя, то документальная ИПС выдаст документы как о ядерных, так и о химических реакторах. Документы по какому-то из этих двух видов реакторов могут не соответствовать информационному запросу. Для предотвращения выдачи нерелевантных документов необходимо устранение полисемии ключевых слов, используемых для координатного индексирования документов и информационных запросов.

5. *Омонимия ключевых слов.* При использовании для координатного индексирования ключевых слов-омонимов вероятны такие же ситуации, какие возникают из-за полисемии ключевых слов (см. пункт 4). Это означает, что омонимия ключевых слов может быть источником «поискового шума» и поэтому должна быть устранена.

6. *Необозначенность родо-видовых связей между ключевыми словами.* Это является причиной невыдачи релевантных документов при информационных запросах тематического характера. Допустим, что предмет информационного запроса — это «Галоиды щелочных металлов». Тогда поисковое предписание будет сформулировано следующим образом: *галоиды*, *щелочные металлы*. В ответ на такое поисковое предписание наша гипотетическая ИПС не выдаст ни одного документа, хотя документ 02 соответствует информационному запросу. Для обеспечения выдачи документа 02 при информационных запросах такого типа необходимо, чтобы в словаре ключевых слов были указаны названия родовых классов этих слов. В данном случае нужно, чтобы ключевые слова *бромиды*, *иодиды*, *хлориды*, *натрий*, *калий* были сгруппированы следующим образом:

а) *галоиды*
— *бромиды*
— *иодиды*
— *хлориды* и т. д.

б) *щелочные металлы*
— *калий*
— *натрий* и т. д.

Тогда поисковое предписание можно было бы сформулировать так: *бромиды*, *калий*, *хлориды*, *натрий*, *иодиды* и т. д. Очевидно, что при такой формулировке поискового предписания наша ИПС выдала бы документ 02.

7. *Ложные синтагматические связи.* Предмет информационного запроса — «Передача электроэнергии из Шотландии

в Англию», поисковое предписание: *передача*, *электроэнергия*, *Шотландия*, *Англия*. Документальная ИПС выдает документ 08 (Передача электроэнергии из Англии в Шотландию), хотя этот документ не соответствует информационному запросу. В данном случае ИПС выдала нерелевантный документ, хотя имело место точное совпадение поискового предписания с поисковым образом документа 08. Отсюда следует, что для предотвращения выдачи нерелевантных документов необходимо, чтобы ключевые слова поисковых образов документов и ключевые слова поисковых предписаний можно было связывать более сильными синтагматическими связями, чем простая координация.

8. *Невыдача документов, близких по смыслу информационному запросу*. Допустим, что поисковый образ документа 11 состоит из ключевых слов $A$, $B$ и $E$, а поисковое предписание — из ключевых слов $A$, $B$, $E$, $F$ и $K$. Тогда наша ИПС не выдаст документа 11, хотя он может представлять определенный интерес для потребителя. «Молчание» ИПС объясняется тем, что в ней не предусмотрена выдача документов, поисковые образы которых содержат не все ключевые слова, образующие поисковое предписание. В связи с этим необходимо отметить, что непременной выдачи документа 11 не обеспечивает и сокращение числа ключевых слов в поисковом предписании. Если из поискового предписания ($A$, $B$, $E$, $F$, $K$) по одному исключать ключевые слова и по каждому варианту поискового предписания производить информационный поиск, то для получения нужной формулировки поискового предписания может потребоваться до $C_5^4+C_5^3=15$ попыток. Очевидно, что такой метод последовательного перебора возможных вариантов поискового предписания приемлем лишь тогда, когда ИПС реализована при помощи ЭЦВМ. Но и в этом случае описанный метод не будет самым эффективным.

Рассмотрение приведенных выше примеров показывает, что для существенного повышения точности и полноты информационного поиска, основанного на применении координатного индексирования, необходимо:

а) устранение синонимии, полисемии и омонимии ключевых слов, используемых в качестве лексических единиц ИПЯ (далее мы будем называть их *дескрипторами*);

б) построение специальных словарей, таблиц или схем, в которых бы эксплицитно были выражены наиболее существенные парадигматические связи между дескрипторами;

в) разработка для дескрипторного ИПЯ такого синтаксиса, который бы позволял использовать при построении поисковых образов документов и поисковых предписаний не только простую координацию дескрипторов, но и более сильные синтагматические связи.

Задача (а) относится к области семантики, т. е. к аспекту отношений слов к предметам и явлениям, которые они обозначают,

а задача (б) — к области отношений между предметами и явлениями, обозначаемыми словами. Совокупность методов и средств, применяемых для решения этих двух задач, называется *контролем за словарным составом ИПЯ* (vocabulary control). Благодаря введению такого контроля обеспечивается:

— использование одинаковых ключевых слов для координатного индексирования одинаковых по смысловому содержанию документов и информационных запросов (это является одним из важнейших условий достаточно высокой точности и полноты информационного поиска);

— достаточно большая полнота выдачи при поиске документов по тематическим запросам;

— возможность эффективного поиска документов, смысловое содержание которых лишь частично соответствует информационному запросу.

Контроль за ключевыми словами, которые используются для координатного индексирования документов и информационных запросов, может иметь разные степени. При нулевом или минимальном контроле предметизатор выбирает ключевые слова для координатного индексирования документа или информационного запроса непосредственно из текста документа без учета того, какие ключевые слова уже использовались ранее для индексирования таких же или близких по смыслу документов и информационных запросов. Предметизатор не устраняет синонимии, полисемии и омонимии используемых им ключевых слов и даже не приводит их грамматические формы к нормальному виду. Но нулевой контроль — это скорей теоретическая возможность, чем регулярный прием. Если он и применяется в реальных ИПС, то лишь при координатном индексировании документов, а не информационных запросов. Индексирование информационных запросов в этом случае должно производиться весьма тщательно и с избыточностью, необходимой для нейтрализации отрицательных явлений, которые порождаются отсутствием словарного контроля при индексировании документов.

При полном контроле за словарным составом ИПЯ предметизатору разрешено использовать для координатного индексирования документов и информационных запросов лишь дескрипторы, т. е. такие ключевые слова, которые содержатся в некотором нормативном списке. В таком списке или словаре полностью устранены синонимия, полисемия и омонимия ключевых слов, а также обозначены определенные парадигматические связи между ними. Наилучшие результаты координатного индексирования документов и информационных запросов по такому словарю может обеспечить лишь предметизатор, хорошо знающий структуру этого словаря, в достаточной мере владеющий методикой координатного индексирования и непременно являющийся специалистом в той отрасли науки или техники, к которой относятся индексируемые

тексты. Если у предметизатора-специалиста нет необходимых знаний в области координатного индексирования, то целесообразно введение контрольного лексикографического редактирования составленных им поисковых образов документов и поисковых предписаний.

В связи с этим необходимо отметить, что еще не вполне ясно, на каком этапе работы ИПС выгодней производить словарный контроль — при вводе документов в систему (т. е. при индексировании документов) или при формулировании поисковых предписаний. Например, Г. П. Лун высказывался за минимальный словарный контроль при вводе документов в ИПС. Он писал: «Чрезмерное редактирование явно увеличивает вероятность влияния (bias) сегодняшних интересов, опыта и точек зрения. Вследствие этого полезность системы уменьшится, если изменятся задачи (emphasis) и интересы. Поэтому кажется очевидным, что чем меньше информация систематизируется и сокращается при вводе, тем больше она будет поддаваться динамической интерпретации при ее выводе» [11]. Если согласиться с такой точкой зрения (а она является достаточно обоснованной), то разработанное Дж. Костелло глубокое координатное индексирование, о котором говорилось ранее, явно будет непригодно для ИПС, рассчитанных на большие поисковые массивы и более или менее длительные сроки эксплуатации.

Основные методы и средства, применяемые для контроля за словарным составом дескрипторного языка, а также для выражения синтагматических отношений между дескрипторами в поисковых образах документов и в поисковых предписаниях, будут подробно рассмотрены в следующих разделах данной главы. Для этого представляется целесообразным несколько уточнить термины «дескриптор» и «дескрипторный язык».

*Дескрипторы* — это предназначенные для координатного индексирования документов и информационных запросов нормативные ключевые слова, которые по определенным правилам отобраны из основного словарного состава того или иного естественного языка и у которых искусственно (при помощи соответствующих отсылок и помет) устранены синонимия, полисемия и омонимия [12].

*Дескрипторным языком* называется специальный ИПЯ, словарный состав которого состоит из дескрипторов, а грамматика — по крайней мере из способа построения поисковых образов документов и поисковых предписаний путем координации соответствующих дескрипторов.

Сейчас уже невозможно установить, кому принадлежит идея координатного индексирования. На конференции по проблемам

---

[11] H. P. Luhn. A statistical approach to mechanized encoding and searching of literary information. «IBM Journal of Research and Development», 1957, v. 1, № 4, стр. 310.

[12] Более строгое определение понятия «дескриптор» дано на стр. 404.

применения классификаций для информационного поиска (Великобритания, 1957 г.) Б. Виккери отмечал, что координатное индексирование (с инвертированной организацией картотеки), по-видимому, применялось еще в Шумере около трех тыс. лет назад. Он сообщил, что найдены глиняные таблички, каждая из которых отведена для какого-либо симптома болезни. На табличке перечислялись названия всех болезней, для которых характерен этот симптом. Очевидно, что если по обнаруженным у больного симптомам подобрать соответствующие таблички, а затем определить, название какой болезни содержится на каждой из них, то это и будет вероятное заболевание, которым страдает данный больной. Таким образом, здесь мы имеем дело с прообразом современных диагностических машин [13].

Несомненно, что принцип координатного индексирования был известен и американскому орнитологу Г. Тейлору, который в 1915 г. получил патент на суперпозиционные (просветные) перфокарты [14]. В геологии этот принцип (в сочетании с суперпозиционными перфокартами) применяется с 1920 г. в системах для идентификации минералов. В 1923 г. француз А. Либé предложил использовать координатное индексирование (также в сочетании с суперпозиционными перфокартами) для поиска в картотеке персонала лиц, обладающих заданным сочетанием характеристик [15].

Первым (или, по крайней мере, одним из первых), кто применил метод координатного индексирования для информационного поиска, по-видимому, был У. Баттен, когда в 1939 г. создавал для английского концерна «Империал кемикл индастриз, лтд» поисковую систему для патентов. Эта система была реализована на суперпозиционных перфокартах. У. Баттен пишет, что идея координатного индексирования у него возникла тогда, «когда он тщательно изучал два списка патентов, один из которых относился к предмету $A$, а другой — к предмету $B$, с целью выбрать из них патенты, относящиеся к сочетанию предметов $A$ и $B$» [16].

В 1947 г. американский математик К. Муэрс разработал и запатентовал систему механизированного поиска документов, основанную на использовании идеи координатного индексирования [17].

---

[13] Proceedings of the International study conference on classification for information retrieval. (Dorking, England, 13th—17th May 1957). London, Aslib, 1957, стр. 106.

[14] H. Taylor. Selective device. Патент США № 1165465; заявлен 14.IX1915 г., выдан 28.XII 1915 г.

[15] H. Liber. Systeme et appareil de classement par feuilles mobiles. Франц. патент № 565475; заявлен 7.XI 1923 г., выдан 28.I 1924 г.

[16] W. E. Batten. Specialized files for patent searching. — В кн.: R. S. Casey and J. W. Perry (Eds). Punched cards. Their applications to science and industry. New York, Reinhold, 1951, стр. 170.

[17] New filing system developed for special collections. «Library Journal», 1948, v. 73, № 10, стр. 794—797; C. N. Mooers. Improvements in or relating to encoding information on punched record cards or other equivalent media; Англ. патент № 681902; заявлен 3.IX 1948 г.

Созданную им ИПС К. Муэрс назвал Системой зато-кодирования (по имени организованной и возглавляемой им консультативной фирмы «Затор ко.»)[18]. Для реализации этой ИПС были созданы особые карты с вырезками вдоль краев, названные «зато-картами», а также специальный селектор. Запись характеристик на таких картах производилась по суперпозиционному коду (с использованием случайных чисел), который был разработан К. Муэрсом и назван «зато-кодированием». Описание зато-карт и общих принципов суперпозиционного кодирования дано в главе 7.

Необходимо отметить, что именно К. Муэрс в 1950 г. ввел для обозначения лексических единиц ИПЯ, применяемого в его поисковой системе, термин «дескриптор». Им были введены также термины «информационный поиск» (information retrieval), «информационно-поисковая система» (information retrieval system), «информационно-поисковый язык» (retrieval language), «дескрипторный словарь» (descriptor dictionary), «поисковый образ» (tally) и другие термины, которые в настоящее время широко используются в области информационного поиска[19].

В одной из своих ранних работ К. Муэрс следующим образом излагал сущность метода индексирования документов, который был применен в созданной им ИПС: «Предмет каждого документа или другого элемента информации характеризуется или описывается с помощью некоторого множества «дескрипторов», взятых из формального словаря таких терминов. В качестве грубого приближения к тому, что здесь имеется в виду, на память приходит «список предметных заголовков» (subject heading list)[20]. И далее: «Каждый документ или элемент информации характеризуется некоторым множеством дескрипторов, взятых из словаря дескрипторов. Каждый дескриптор этого множества применим или в каком-то смысле истинен в отношении информационного содержания элемента информации. При таком типе описания (characterization) дескрипторы функционируют независимо друг от друга. Тот факт, что данное множество состоит из нескольких дескрипторов, может означать, что в исходном документе они образуют какое-то взаимодействующее сочетание; это вполне может также означать, что они относятся к разным самостоятельным идеям, рассеянным в [тексте] документа»[21].

По К. Муэрсу, дескриптор — это слово или словосочетание, обозначающее какое-либо более или менее широкое понятие. При этом не требуется, чтобы значение дескриптора строго соответствовало тому значению, в котором обычно употребляется

---

[18] Фирма была создана в 1948 г. (г. Кембридж, штат Массачусетс); с 1962 г. она называется «Рокфорд рисерч институт инк.».

[19] C. N. Mooers. Zatocoding applied to mechanical organization of knowledge. «American Documentation», 1951, v. 2, № 1, стр. 20—32.

[20] Там же, стр. 25.

[21] Там же, стр. 26.

единичные понятия (unit concepts). Все унитермы имеют одинаковый иерархический ранг: ни один из них в логическом смысле не стоит выше или ниже какого-либо другого (как в классификации) и ни один не используется в заранее установленном сочетании с каким-либо другим (как в традиционных предметных заголовках)» [25].

Словосочетание может быть унитермом лишь тогда, когда ни одно из образующих его слов не входит в состав другого словосочетания, используемого в данном ИПЯ в качестве лексической единицы. Например, словосочетание «суперпозиционные перфокарты» выступают в функции унитерма, если при координатном индексировании мы не используем словосочетаний «суперпозиционные коды», «щелевые перфокарты» и т. д. Простые унитермы М. Таубе называет *свободными*, а составные — *несвободными* или *связанными* (bound). Связанные унитермы включаются в словарный состав ИПЯ как в прямом, так и в инвертированном виде. Например: «Суперпозиционные перфокарты» и «Перфокарты, суперпозиционные». Исключение из этого правила составляют словосочетания, обозначающие имена собственные: они даются только в прямом виде [26].

Такой принцип построения словарного состава ИПЯ позволяет сократить его объем до минимума. В словнике алфавитно-предметного каталога Отдела технической информации Библиотеки конгресса США в 1952 г. насчитывалось около 50 тыс. предметных заголовков. Преобразование этого словника в список унитермов показало, что в данном словнике использовалось всего лишь около трех тыс. слов [27]. В связи с этим полезно напомнить, что в любом естественном языке 1000 самых употребительных слов составляют 80% всех текстов.

Мы видим, что в системе унитермов М. Таубе в какой-то степени используется (сознательно или бессознательно) принцип [illegible]льных словарей» Б. Рассела. Термин «минимальный сло- Рассел определяет следующим образом: «Я называю минимальным», если в нем не содержится ни одного [illegible]орое можно было бы определить с помощью других слов [illegible]аря» [28]. Это определение может породить опасную иллю- построение минимального словаря — сравнительно про- адача. На самом же деле построить такой словарь очень

[25] J. C. Costello. Jr. Uniterm indexing principles, problems and solutions. «American Documentation», 1961, v. 12, № 1, стр. 21.

[26] M. Taube, C. D. Gull and I. S. Wachel. Unit terms in coordinate indexing. «American Documentation», 1952, v. 3, № 4, стр. 213—218.

[27] R. B. Thomas and C. D. Gull. The choice of terms for a Uniterm coordinate index of scientific and technical reports. — В кн.: M. Taube et al. Studies in coordinate indexing, v. 1, Washington, Documentation, Inc., 1953, стр. 47—55.

[28] B. Russel. Human knowledge: its scope and limits. London, Allen and Unwin, 1948, стр. 94, 274—283.

трудно, так как для этого необходимо решить множество сложных семантических проблем.

Для технической реализации информационно-поисковых систем, основанных на использовании системы унитермов, М. Таубе разработал специальную карточку, которая была названа им «унитерм-картой». Унитерм-карта представляет собой прямоугольный кусок плотной бумаги (обычно — формата 75×125 *мм* или 203×125 *мм*), который снабжен специальной сеткой, состоящей из одной горизонтальной и 10 вертикальных граф (рис. 45). В горизонтальной графе, занимающей верхнюю часть карточки, записывается унитерм. Такие карты заводятся на все унитермы, используемые в системе. Если унитерм состоит из двух и более слов, то унитермы заводятся на все возможные перестановки из этих слов. Унитерм-карты расстанавливаются в алфавитном порядке унитермов и хранятся в обычных каталожных ящиках. В вертикальных графах записываются адресные шифры (номера) документов, в поисковые образы которых входит унитерм, указанный в верхней графе унитерм-карты.

Особенность записи номеров документов на унитерм-картах состоит в том, что она производится по последней цифре номера. Например, номер 127 записывается в графе 7, номер 239 — в графе 9, номер 270 — в графе 0 и т. д. Поскольку документы получают номера и заносятся на унитерм-карты по мере поступления, в графах унитерм-карт они будут располагаться в возрастающем порядке. Такая система записи значительно облегчает процедуру визуального выявления номеров документов, которые одновременно содержатся в каждой из двух и более сравниваемых унитерм-карт. Эта процедура эквивалентна отысканию номеров документов, основное смысловое содержание которых выражается совокупностью (логическим произведением) унитермов, обозначающих сравниваемые унитерм-карты.

Предположим, что нам необходимо ввести в ИПС статью, в которой рассматриваются методы защиты от коррозии лопаток газовых турбин. Для координатного индексирования этой статьи будут выбраны унитермы «защита», «коррозия», «газ», «турбины» и «лопатки», которые в совокупности и образуют ее поисковый образ. Статье присваивается очередной адресный шифр 2005. Затем из картотеки извлекаются унитерм-карты, обозначенные словами «защита», «коррозия», «газ», «турбины» и «лопатки», и на каждой из этих унитерм-карт в графе 5 проставляется номер 2005. После этого унитерм-карты расстанавливаются в картотеке на свои места. Такая процедура повторяется при вводе в ИПС каждого нового документа.

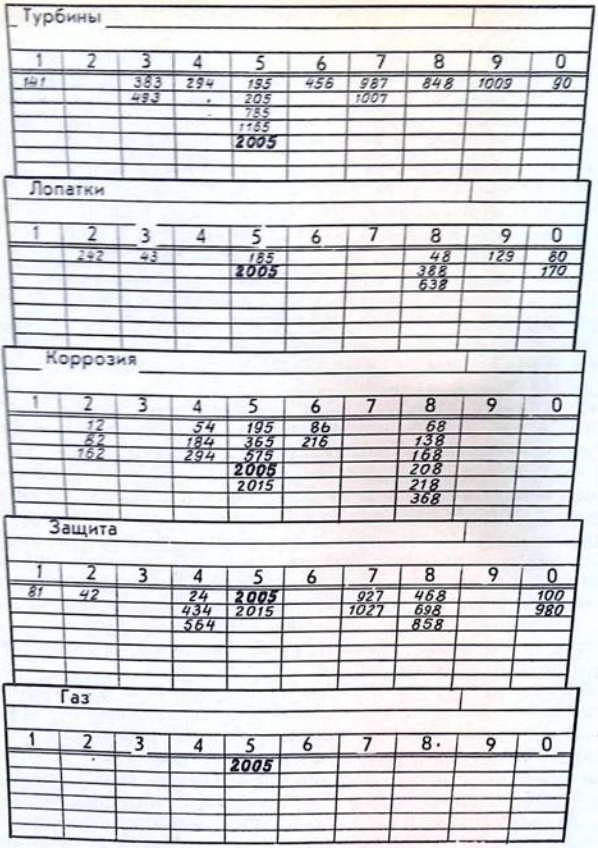
Информационный поиск в ИПС, созданной на базе картотеки унитерм-карт, производится следующим образом. Допустим, что в ИПС поступил информационный запрос о публикациях, посвященных лопаткам газовых турбин. На основании этого запроса

1) Ввод в ИПС статьи, в которой рассматриваются вопросы защиты от коррозии лопаток газовых турбин (документ 2005).

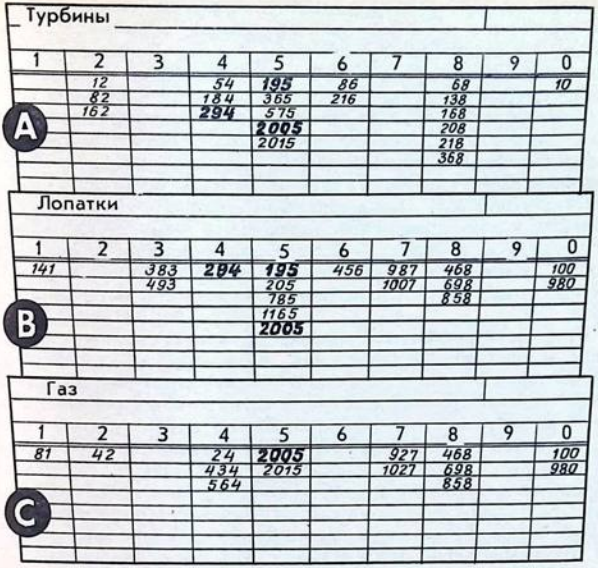
Рис. 45. Принцип действия ИПС, реализованной на унитерм-картах

2) Поиск в ИПС документов, темой которых являются лопатки газовых турбин. В унитерм-картах А и В совпадают три номера — 195, 294 и 2005. Это означает, что в документах под такими номерами говорится о лопатках турбин. Из этих трех номеров в унитерм-карте С содержится лишь номер 2005. Именно документ под таким номером и отвечает на информационный запрос.

Рис. 45 (*продолжение*)

документов [35]. В другом случае, когда создавалась система унитермов по психологии, появление новых унитермов почти прекратилось после индексирования около четырех тыс. документов [36].

При создании документальной ИПС в Центре технических данных фирмы «Дженерал электрик ко.» (США) число унитермов после индексирования более 150 тыс. научно-технических отчетов составило всего около трех тыс.[37] Опыт создания этой системы показал, что рост числа унитермов в ее словаре описывается асимптотической функцией и что для одной отрасли науки или техники

Рис. 46. Рост числа унитермов в зависимости от числа заиндексированных документов

A — предел роста

предел роста (асимптота) — около трех тыс. унитермов. Для нескольких отраслей науки и техники, исключая органическую химию, рост словаря почти прекращается при достижении им объема примерно семи тыс. унитермов. Если же в число этих отраслей включить и органическую химию, то «насыщение» словаря наступает при накоплении от 12 до 15 тыс. унитермов [38]. Что же касается темпа роста числа унитермов в словаре, то он находится в прямой зависимости от глубины и широты индексирования, т. е. от среднего числа унитермов в поисковом образе каждого документа.

По-видимому, асимптотический рост числа унитермов в зависимости от числа заиндексированных документов может быть описан формулой

$$y = A(1 - e^{-kx}),$$

где $A$ — предел, к которому стремится общее число унитермов в словаре, $e$ — основание натуральных логарифмов ($e = 2{,}718\ldots$), $k$ — константа роста, $x$ — число заиндексированных документов.

---

[35] J. A. Sanford and F. R. Theriault. Problems in the application of uniterm coordinate indexing. «College and Research Libraries», 1956, v. 17, № 1, стр. 20.

[36] P. L. Broadhurst. Coordinate indexing: a bibliographic aid. «American Psychologist», 1962, v. 17, № 3, стр. 140.

[37] R. L. Francisco. Use of the Uniterm coordinate indexing system in a large industrial concern. «Special Libraries», 1956, v. 47, № 3, стр. 118.

[38] Там же, стр. 119.

Величина $A$ зависит от отрасли науки или техники, к которой относятся индексируемые документы, и может иметь значения примерно от $3 \cdot 10^2$ до $1{,}5 \cdot 10^4$. Величина коэффициента $k$ зависит от средней глубины индексирования документов. График асимптотической функции приведен на рис. 46.

В заключение данного раздела необходимо еще раз остановиться на действительных и мнимых преимуществах дескрипторных языков перед библиотечно-библиографическими и алфавитно-предметными классификациями. Основное достоинство дескрипторных языков заключается в том, что такие языки позволяют осуществлять информационный поиск по любому, заранее не заданному сочетанию характеристик, т. е. обладают большей «семантической силой» и гибкостью, чем библиотечно-библиографические и алфавитно-предметные классификации. В связи с этим следует отметить, что иногда дескрипторным языкам приписываются и такие достоинства, которыми эти языки на самом деле не обладают.

Например, встречаются утверждения, что одной из основных причин все более широкого применения дескрипторных языков является то, что при их использовании достигается большая экономия места, занимаемого документальной ИПС. Такие утверждения не выдерживают никакой критики. Если систематический или алфавитно-предметный каталог какой-либо крупной библиотеки удовлетворительно выполняет свои функции, то в настоящее время едва ли может возникнуть вопрос о его замене поисковой системой с дескрипторным языком только потому, что этот каталог занимает слишком много места.

Вообще сокращение физического объема, занимаемого документальной ИПС, — за исключением весьма немногочисленных особых случаев — сейчас не относится к числу первоочередных задач. Более того, такая задача пока не является важнейшей даже для самых больших библиотек мира, в фондах которых насчитываются многие миллионы документов. Сказанное не следует понимать в том смысле, что нам не нужны специальные устройства, позволяющие хранить в виде микрокопий огромные массивы документов и по требованию быстро выдавать увеличенные отпечатки тех или иных документов. Наоборот, такие устройства очень нужны, но их создание не решает и не может решить ни одной из основных проблем документального информационного поиска.

В качестве одного из основных преимуществ координатного индексирования перед традиционными системами иногда называют то, что при его применении производится более глубокое индексирование документов. Например, декан библиотечного училища Мэрилендского университета (США) П. Вассерман в своей книге «Библиотекарь и машина» пишет: «Для систематических и алфавитно-предметных систем типичны небольшая глубина индексирования и, следовательно, экономичность процесса ввода. Но они требуют больше затрат времени и кропотливой работы на

выводе, когда производится информационный поиск, который поэтому является более дорогостоящей процедурой. При использовании таких систем обычны задержки в обнаружении требуемого материала, а поиск публикаций может быть затруднительным или неадекватным. В качестве типичной черты координатных указателей, основанных на использовании «унитермов», «дескрипторов», «ключевых слов» или «аспектов», можно назвать бóльшую глубину индексирования. На этапе ввода каждому документу приписывается больше терминов. Для информационного поиска производится координация терминов, описывающих искомую информацию. Таким образом, при координатном индексировании ввод является сравнительно более точным и подробным, а следовательно, и более дорогостоящим, тогда как вывод должен быть более эффективным, быстрым и экономичным» [39].

В целом это соответствует *сложившейся практике* (sic!). Однако необходимо подчеркнуть, что небольшая глубина индексирования, практикуемая при использовании традиционных библиотечно-библиографических систем, обусловлена отнюдь не имманентными свойствами этих систем, а скорей традиционными средствами их реализации. При значительном увеличении глубины индексирования карточные каталоги становятся неудобными для использования. Едва ли кто-нибудь рискнет утверждать, что иерархическая или алфавитно-предметная классификация не позволяет заиндексировать документ с любой степенью глубины. Поэтому, когда мы сравниваем «семантическую силу» информационно-поисковых языков, то необходимо исходить главным образом из внутренних свойств этих систем, а не из того, что одни из них по тем или иным причинам функционируют в более благоприятных, а другие — в менее благоприятных условиях.

К числу основных достоинств дескрипторных языков относят также и то, что они позволяют механизировать и автоматизировать информационный поиск. При этом молчаливо предполагается, что механизация и автоматизация информационного поиска есть благо при всех условиях. Такая точка зрения, как уже отмечалось ранее, глубоко ошибочна. К сказанному нужно еще добавить, что при использовании дескрипторных языков механизация и автоматизация информационного поиска в какой-то мере является суровой необходимостью, ибо ручные методы поиска по сочетаниям дескрипторов, как правило, весьма не эффективны. Исключение составляют лишь ИПС дескрипторного типа, реализованные при помощи унитерм-карт или выполненные в виде специальных изданий (имеются в виду спаренные альбомы унитерм-карт, указатели табличного типа и т. д.).

---

[39] P. Wasserman. The librarian and the machine. Observations on the applications of machines in administration of college and university libraries. Detroit, Gale Research Co., 1965, стр. 57.

В мае 1963 г. Центр по обработке научной информации Евратома опубликовал результаты анкетного опроса 218 учреждений и организаций разных стран мира об основных характеристиках созданных ими ИПС [40]. Опрос был проведен в 1961 г. Ответы поступили из 18 стран, главным образом — из Великобритании, Италии, Нидерландов, Франции, ФРГ и других западноевропейских стран. США этим опросом были охвачены в минимальной степени.

Сведения о типе применяемого ИПЯ были сообщены лишь в 121 ответе. Анализ этих сведений показывает, что дескрипторные языки применяются в 70 из 121 системы (57%). Подавляющее большинство этих систем (57) механизировано или полумеханизировано, т. е. реализовано при помощи ЭЦВМ, счетно-перфорационных машин или перфокарт ручного обращения. Библиотечно-библиографические и алфавитно-предметные классификации применяются в 52 системах (43%), причем подавляющее большинство из них (38) реализовано при помощи обычных каталожных карточек. Из этих 52 систем только 4 реализованы при помощи счетно-перфорационных машин и ни одной — при помощи ЭЦВМ; 60% ИПС с классификационным индексированием было создано в период с 1957 по 1961 г. Сводные данные о распределении 121 ИПС по типам применяемого в них ИПЯ и техническим средствам их реализации приведены в табл. 23

Таблица 23

**Распределение информационно-поисковых систем по типам применяемого в них ИПЯ и средствам их реализации**

| Тип ИПЯ | Технические средства реализации ИПС | | | | Всего |
|---|---|---|---|---|---|
| | ЭЦВМ | Счетно-перфорационные машины | Перфокарты ручного обращения | Каталожные карточки | |
| Дескрипторные языки | 16 | 28 | 13 | 13 | 70 |
| УДК . . . . . | — | 1 | 4 | 19 | 24 |
| Другие виды классификаций . . . . . | — | 3 | 6 | 18 | 27 |
| Всего: | 16 | 32 | 23 | 50 | 121 |

Столь незначительное применение библиотечно-библиографических классификаций в механизированных и автоматизированных ИПС объясняется не тем, что классификации обладают некими

[40] H. H. Bernstein and S. Gabbai. Enquiry on non-conventional documentation systems in use. Final report sponsored by FIS/MSR Committee, EUR 331e. Joint Nuclear Research Center — CETIS, 1963.

внутренне присущими им свойствами, которые препятствуют механизации процедуры сопоставления классификационных индексов документов с индексами информационного запроса. Очевидно, что механизация такой процедуры легко осуществима. Но если картотека упорядочена по классификационным индексам (например, по индексам УДК), то человек отыскивает в ней нужные ему карточки настолько быстро, что применение для этой цели машин чаще всего бывает просто ненужным. Именно поэтому большинство ИПС, в которых для индексирования документов применяются УДК и другие библиотечно-библиографические классификации, реализуется при помощи обычных каталожных карточек.

Особенно широкое распространение дескрипторные языки получили в США. В октябре 1962 г. Национальный научный фонд США опубликовал обзор, в котором описано 87 действующих ИПС, основанных на использовании координатного индексирования [41]. Из них 59 являются механизированными (22 реализованы с помощью ЭЦВМ) и 19 — полумеханизированными. При помощи каталожных карточек реализованы только 2 системы и при помощи унитерм-карт — 7. В обзор включены лишь три ИПС, находящиеся за пределами США, в том числе одна ИПС на каталожных карточках, созданная в Канаде в 1935 г.

Обсуждение — даже в самом общем виде — проблемы механизации и автоматизации информационного поиска будет неполным, если не сделать попытки оценить роль, которую призваны сыграть в этой области ЭЦВМ. Исключительно успешное применение вычислительных машин в самых разнообразных областях человеческой деятельности породило у многих людей почти мистическую веру в их всемогущество и непогрешимость. Однако действительные возможности вычислительных машин далеко не безграничны. Основные преимущества вычислительных машин перед мозгом человека заключаются в следующем:

1) вычислительная машина обеспечивает полное единообразие выполнения операций. При использовании вычислительной машины документ всегда будет заиндексирован одинаково, тогда как один и тот же человек при разных обстоятельствах или два разных человека могут заиндексировать этот документ по-разному. Э. М. Маккормик писал по этому поводу: «Вычислительная машина о б ъ е к т и в н а (objective) при выполнении операций. Она будет повторять то, что ей поручено. Если о каком-либо документе дано определенное количество информации, то вычислительная машина всегда будет т о ч н о (in a precise manner) индексировать этот документ, но насколько п р а в и л ь н о (accurately) — это уже другой вопрос» [42];

[41] Nonconventional technical information systems in current use, № 3. NSF-62-34. Washington, National Science Foundation, 1962.

[42] E. M. McCormick. Why computers? В кн.: «Machine indexing: progress and problems». Papers presented at the 3rd Institute on Information

2) вычислительная машина обычно весьма надежна в работе и допускает минимальное число ошибок. Этим она выгодно отличается от человека, который обычно делает значительно больше ошибок, причем с увеличением его усталости число ошибок быстро возрастает;

3) вычислительная машина обладает большим быстродействием;

4) стоимость выполнения вычислительной машиной некоторых операций может быть значительно ниже стоимости выполнения таких же операций человеком.

Все эти несомненные достоинства вычислительной машины могут быть использованы лишь при условии, что разработан алгоритм и составлена программа решения соответствующей задачи. Более ста лет назад А. Лавлейс — дочь Байрона и, в сущности, первый программист цифровой вычислительной машины — сказала по поводу универсальной вычислительной машины, задуманной Чарльзом Бэббеджем в 1833 г.: «Аналитическая машина не претендует на то, чтобы создавать что-то *новое*. Она может делать лишь то, выполнение чего *мы знаем, как ей предписать*» [43]. Это высказывание в целом справедливо и в наши дни, хотя развитие эвристического программирования и самопрограммирования может радикально изменить существующее положение.

Для разработки алгоритма любой процедуры необходимо понимание внутреннего механизма ее выполнения человеком. В настоящее время мы еще не знаем, как человек производит реферирование, аннотирование и индексирование документов, при котором требуется определение основного смыслового содержания этих документов. Поэтому алгоритмизация таких процедур пока невозможна.

Вычислительная машина может значительно быстрей и точней, чем человек, сравнивать текст информационного запроса с полными текстами документов. Но она это делает, основываясь лишь на статистических свойствах этих текстов и не касаясь их смыслового содержания. Пока еще не доказано, что смысловая эквивалентность двух текстов на естественном языке может быть установлена статистическими методами без значительно более сложного смыслового анализа этих текстов, требующего использования внеязыковой информации, т. е. такой, которая непосредственно в обрабатываемых текстах не содержится.

Из сказанного выше было бы неправильно сделать вывод, что вычислительные машины никогда не смогут решать такие задачи, как аннотирование, реферирование и индексирование. Возмож-

---

and Retrieval, February 13—17, 1961. 2nd print. 1963, Center for Technology and Administration, The American University, стр. 220.

[43] Цит. по: B. V. Bowden (Ed.). Faster than thought. A symposium on digital computing machines. London, Pitman, 1953, стр. 398.

ности вычислительных машин изучены еще сравнительно мало. Нет сомнения в том, что когда-нибудь эти машины будут успешно реферировать, аннотировать и индексировать документы. Но для этого нужны глубокие исследования в области психологии мышления, лингвистики, математической логики, семиотики и т. д. Что же касается сегодняшнего дня, то наиболее перспективным нам представляется путь отыскания оптимальных форм взаимодействия человека и машины, при котором их преимущества использовались бы в наиболее полной мере и наилучшим образом дополняли друг друга. Вот что писал о таком взаимодействии Н. Винер: «Вряд ли можно считать, что мозг в сравнении с современными вычислительными машинами не имеет определенных преимуществ, связанных с его огромным функциональным диапазоном, несоизмеримо большим, чем можно было бы ожидать, учитывая его физические размеры.

Главное из этих преимуществ — по-видимому, способность мозга оперировать с нечетко очерченными понятиями. В таких случаях вычислительные машины, по крайней мере в настоящее время, почти не способны к самопрограммированию. Между тем наш мозг свободно воспринимает стихи, романы, картины, содержание которых любая вычислительная машина должна была бы отбросить, как нечто аморфное.

Отдайте же человеку — человеческое, а вычислительной машине — машинное. В этом и должна, по-видимому, заключаться разумная линия поведения при организации совместных действий людей и машин. Линия эта в равной мере далека и от устремлений машинопоклонников, и от воззрений тех, кто во всяком использовании механических помощников в умственной деятельности усматривает кощунство и принижение человека» [44].

## СИНОНИМИЯ КЛЮЧЕВЫХ СЛОВ И ЕЕ УСТРАНЕНИЕ

В любом естественном языке один предмет может иметь не одно, а два и более различных имен (знаков). Другими словами, одно и то же «означаемое» может иметь два и более «означающих» [45]. Это явление естественного языка называется синонимией, а «слова, способные в том же контексте или контекстах, близких по смыслу, заменять друг друга без ощущения заметного различия по смыслу, носят название синонимов (*греч.* synónymos —

---

[44] Н. Винер. Творец и робот. Обсуждение некоторых проблем, в которых кибернетика сталкивается с религией. М., «Прогресс», 1966, стр. 82—83.

[45] В языкознании означаемым называется предмет мысли, внутренняя (семантическая) сторона языковой единицы, а означающим — внешняя (материальная) сторона этой единицы.

«одноименный», «со-имя»)»[46]. Синонимами являются, например, следующие слова: *лингвистика — языкознание — языковедение*; *опыт — эксперимент*; *абстрактный — отвлеченный — умозрительный* и т. д. Необходимо отметить, что синонимами могут быть не только отдельные слова, но и такие выражения, как *транзистор* и *полупроводниковый триод*; *квадрат* и *равносторонний прямоугольник*; *бронза* и *сплав меди с оловом* и т. д.

Абсолютных (тождественных) синонимов[47] в естественных языках очень мало. Некоторые философы и языковеды (например, Д'Аламбер, Л. Блумфилд и др.) считают, что в языке вообще не может быть двух слов, которые можно было бы считать абсолютными синонимами, т. е. такими словами, которые во всех случаях могли бы заменять друг друга без изменения целого. Подавляющее большинство синонимов, имея концептуально одинаковое значение, обладает различными количественными и качественными оттенками.

Синонимия может быть лексической и грамматической. В основе лексической синонимии лежит употребление двух и более разнокорневых слов для обозначения одного предмета или понятия, а в основе грамматической синонимии — употребление для этой же цели слов, которые имеют один общий корень, но разный морфологический состав и входят в разные грамматические категории. Слова *алфавит* и *азбука* являются лексическими синонимами, а слова *студенты* и *студенчество* — синонимами грамматическими (иногда такие синонимы называют однокорневыми). Для построения тезаурусов основное значение имеет лексическая синонимия.

Лексические синонимы можно разделить на три группы: контекстуальные, тематические и подлинные. Контекстуальный (или фразеологический) синоним — это слово, которым можно заменить некоторое другое слово только в данном конкретном контексте. Очевидно, что контекстуальные синонимы не могут попасть в список ключевых слов, содержательно выбранных из текста документа для координатного индексирования этого документа. Поэтому такие синонимы в данном случае не рассматриваются[48]. Тематические синонимы возникают в том случае когда без какой-либо дифференциации употребляются слова, связанные с одной широкой темой. Например, тематическими синонимами слова (*пере*)*движение* могут быть следующие слова: *ходьба*, *бег*, *езда*, *скачка*, *полет*, *плавание* и т. д. Тематический ряд образуют также род и его видовые понятия: *публикация — заметка*, *статья*, *отчет*, *монография*, *учебник*, *справочник* и т. д.

---

[46] Л. А. Булаховский. Введение в языкознание, ч. 2. М., Учпедгиз, 1953, стр. 39.

[47] Их называют еще лексическими дублетами.

[48] Контекстуальные синонимы необходимо учитывать при машинном индексировании текстов.

Слова-синонимы образуют незамкнутые синонимические ряды (как двучленные, так и многочленные). В синонимический ряд могут быть включены лишь слова, принадлежащие к одной части речи. Из числа слов, образующих синонимический ряд, выделяется слово, определяющее характер всего синонимического ряда. Такое слово называется доминантой. В качестве доминанты всегда выступает слово, наиболее нейтральное в стилистическом отношении, не содержащее какой-либо оценочной характеристики того, что им называется. По отношению к доминанте все остальные члены синонимического ряда являются зависимыми словами. Каждое слово синонимического ряда должно быть синонимичным не только доминанте, но и всем остальным словам данной группы. Это означает, что для всех без исключения членов синонимического ряда должно быть характерно, по крайней мере, какое-либо одно общее значение. Для того чтобы синонимический ряд был достаточно четким, доминанту необходимо брать только в одном, отграниченном значении.

Необходимо отметить, что из-за явления полисемии у одного и того же слова может быть целый ряд синонимов, которые между собой в синонимических отношениях находиться не будут [49].

Для русского языка (как и для других языков) характерен двучленный или трехчленный синонимический ряд. В русском языке наибольшее число синонимов наблюдается среди прилагательных, а наименьшее — среди существительных.

Ранее уже отмечалось, что в качестве ключевых слов чаще всего выступают термины. Термины — это слова или словосочетания, стремящиеся быть однозначными. Обязательными требованиями к терминам являются строгая определенность значения и устойчивость употребления. В принципе термины не должны иметь синонимов. Однако в результате воздействия тех или иных обстоятельств для названия (обозначения) одного и того же предмета нередко употребляется несколько параллельных терминов, например: *триггер — спусковая схема*; *транзистор — полупроводниковый триод*; *аналоговая вычислительная машина — вычислительная машина непрерывного действия*, *моделирующее устройство* и т. д. Таким образом, синонимия — явление, свойственное не только общелитературному языку, но и специальной научно-технической терминологии.

При координатном индексировании предметизатор испытывает на себе сильное влияние лексики, которую использовал автор индексируемого документа. Но разные авторы для описания одних

---

[49] В. Н. Клюева. Краткий словарь синонимов русского языка. М., Учпедгиз, 1961, стр. 4—8; Н. М. Шанский. Лексикология современного русского языка. М., «Просвещение», 1964, стр. 53—55.

и тех же предметов и явлений часто употребляют разные слова. Поэтому если допустить бесконтрольное употребление синонимичных ключевых слов при координатном индексировании (т. е. выбирать их непосредственно из индексируемого текста), то документам с одинаковым смысловым содержанием (предметом) могут присваиваться поисковые образы, состоящие из разных ключевых слов. Это означает, что для обеспечения исчерпывающей полноты информационного поиска индексатору пришлось бы включать в каждое поисковое предписание все синонимы ключевых слов, описывающих смысловое содержание информационного запроса. Если пойти по такому пути, то поисковое предписание нужно формулировать:

— либо в виде конъюнкции дизъюнкций ключевых слов, описывающих смысловое содержание информационного запроса, и всех без исключения синонимов этих ключевых слов,

— либо в виде дизъюнкции конъюнкций этих же ключевых слов и их синонимов.

Эти конъюнкции образуют подпредписания, представляющие собой различные сочетания, которые можно составить из ключевых слов, выражающих смысловое содержание информационного запроса, или из синонимов этих ключевых слов. Если поисковое предписание составлено из ряда подпредписаний, то выдача на это предписание получается объединением выдач на каждое из составляющих его подпредписаний.

Необходимо отметить, что слово «конъюнкция» в данном случае употребляется для обозначения логического произведения ключевых слов, их синонимов или подпредписаний, о которых говорилось выше. (Ср. употребление соединительного союза «и» в грамматике). Такой вид связи мы ранее условились обозначать символом &.

Соответственно слово «дизъюнкция» следует понимать в том смысле, что вместо ключевого слова может быть подставлен любой его синоним. В грамматике примерно такую же функцию выполняет разделительный союз «или». Дизъюнкцию в указанном выше смысле мы условились обозначать символом $+$.

Пусть $A$, $B$ и $C$ — ключевые слова, выражающие смысловое содержание некоторого информационного запроса. $A'$, $A''$, $B'$ и $C'$ — соответственно синонимы ключевых слов $A$, $B$ и $C$. Тогда поисковое предписание должно быть сформулировано в одном из следующих видов:

$$1)\ (A + A' + A'') \,\&\, (B + B') \,\&\, (C + C');$$

$$2)\ (A \,\&\, B \,\&\, C) + (A' \,\&\, B \,\&\, C) + (A'' \,\&\, B \,\&\, C) + \ \dots\ + (A'' \,\&\, B' \,\&\, C').$$

12 подпредписаний

Для вычисления общего числа правильно построенных подпредписаний такого типа может быть использована следующая формула [50]:

$$k=\prod_{i=1}^{m}(n_i+1),$$

где $k$ — общее число правильно построенных подпредписаний; $m$ — число разных ключевых слов, используемых для выражения смыслового содержания информационного запроса; $n_i$ — число синонимов $i$-того ключевого слова ($i=1, 2, 3, \ldots, m$). По этой формуле нетрудно подсчитать, что при $m=6$, $n_1$, $n_3=2$, $n_2$, $n_4=1$ и $n_5$, $n_6=0$ (по-видимому, такие значения $m$ и $n_i$ можно считать вполне реальными) $k=36$.

Из сказанного выше следует, что если при координатном индексировании допустить бесконтрольное употребление синонимичных ключевых слов, то это значительно усложнит процедуру информационного поиска. Особенно трудно будет составлять правильные поисковые предписания, т. е. переводить информационные запросы на применяемый ИПЯ. Для этого индексатор должен не только знать все ключевые слова и их синонимы, но и активно пользоваться ими. Но это практически невозможно. Поэтому индексатора, осуществляющего перевод информационных запросов на ИПЯ, пришлось бы снабдить предварительно составленным словарем синонимов ключевых слов. Только при этом условии можно было бы рассчитывать, что в подавляющем большинстве случаев информационные запросы будут правильно переводиться на ИПЯ.

Однако предварительное составление словаря синонимов ключевых слов и использование этого словаря при переводе информационных запросов на ИПЯ означает не что иное, как установление контроля над употреблением синонимичных ключевых слов, хотя и не при индексировании документов.

Нетрудно понять, что процедура информационного поиска существенно упростится, если синонимия ключевых слов будет вообще устранена. Тогда при координатном индексировании документам, имеющим одинаковое смысловое содержание, будут присваиваться поисковые образы, состоящие из одних и тех же ключевых слов.

В естественном языке устранить синонимию в принципе невозможно, так как это лишило бы нас возможности описывать, выражать и определять одни явления через другие, объяснять вновь вводимые слова через другие, уже известные нам слова. В сущности, устранение синонимии в естественных языках озна-

[50] Правильно построенным подпредписанием мы называем такое, в которое не входят одновременно ключевое слово и его синонимы (один или несколько).

чало бы ликвидацию науки [51]. Но синонимия может быть устранена в искусственном языке, каким является дескрипторный ИПЯ.

В целях устранения синонимии в списке отобранных для такого ИПЯ ключевых слов необходимо произвести их группировку по смысловому признаку, объединив в отдельные классы (семантические ряды) все эквивалентные по смыслу ключевые слова, и обозначить каждый такой класс доминантой, выбранной из числа слов обозначаемого класса. Тогда все слова, входящие в этот класс, будут определены как синонимы обозначающего этот класс слова-доминанты. Если ключевое слово не имеет синонимов, то оно образует класс, состоящий лишь из одного этого слова.

В элементарных терминах теории множеств процедуру такой группировки можно описать следующим образом. Пусть $T$ — множество всех ключевых слов, отобранных для данного дескрипторного ИПЯ

$$T = \{t_1,\ t_2,\ t_3,\ \ldots,\ t_m\},$$

где $t_1, t_2, t_3, \ldots, t_m$ — отдельные ключевые слова. Очевидно, что множество $T$ может быть разбито на семейство подмножеств $\{T_i\}$. $T_i = \{t_{i1},\ t_{i2},\ t_{i3},\ \ldots,\ t_{ik}\}$, где $t_i \in T$. При таком разбиении $\{T_i\}$ между элементами $t_{i1}, t_{i2}, t_{i3}, \ldots, t_{ik}$ любого подмножества $T_i$ имеет место отношение эквивалентности, т. е. такое отношение, которое обладает свойствами рефлексивности $(t_{i1} \equiv t_{i1})$, симметрии $(t_{i1} \equiv t_{i2} \Rightarrow t_{i2} \equiv t_{i1})$ и транзитивности $(t_{i1} \equiv t_{i2},\ t_{i2} \equiv t_{i3} \Rightarrow t_{i1} \equiv t_{i3})$.

Отношение называется рефлексивным (возвратным) тогда, и только тогда, когда каждый предмет находится в этом отношении к самому себе. Рефлексивными являются отношения равенства и одновременности. Отношение между $a$ и $b$ называется симметричным тогда, и только тогда, когда оно имеет место также между $b$ и $a$. Симметричными являются отношения равенства, сходства, различия, одновременности, некоторые отношения родства и т. д. Отношение называется транзитивным тогда, и только тогда, когда из факта наличия такого отношения между $a$ и $b$, а также $b$ и $c$ вытекает, что оно имеет место и между $a$ и $c$. Транзитивными являются отношения равенства, отношения «больше» и «меньше», отношения «быть раньше», «быть позже», «быть выше», «быть севернее», «находиться под» и т. д. Из приведенных определений следует, что лишь отношения равенства и одновременности являются и рефлексивными, и симметричными, и транзитивными.

Множество вида $T_i$ называется классом эквивалентности. Из каждого такого класса эквивалентности $T_i$ можно было бы выбрать доминанту $t_i \in T_i$, обозначить этой доминантной $t_i$ весь класс эквивалентности $T_i$, формально определив остальные элементы $T_i$ как синонимы $t_i$. Таким путем было бы

[51] Д. П. Горский. Логика. М., Учпедгиз, 1963, стр. 48.

установлено взаимно однозначное соответствие между $t_i \in T_i$ и остальными словами $T_i$ и обеспечена возможность прямой подстановки $t_i$, вместо любого другого элемента $T_i$, т. е. была бы формально устранена синонимия в словарном составе ИПЯ.

Однако отношение эквивалентности в том смысле, в каком оно было определено выше, может иметь место лишь между абсолютными синонимами, которых в естественных языках вообще и в терминологии в особенности очень мало. Поэтому для устранения синонимии в дескрипторных ИПЯ обычно используется принцип условной эквивалентности, по которому ключевые слова объединяются в классы эквивалентности на том основании, что они выражают близкие по содержанию понятия. Смысловая доминанта, выбранная для обозначения какого-либо класса условной эквивалентности и выражающая главное значение всех слов и словосочетаний, которые входят в этот класс, называется *дескриптором*. Однако фактически дескриптором является не имя класса условной эквивалентности, а сам этот класс. В этой связи необходимо отметить, что в литературе термин «дескриптор» обычно употребляется в более широком значении: дескриптором называют всякое ключевое слово, используемое для координатного индексирования документов и информационных запросов.

Дескриптор можно определить так же как лексическую единицу семантического языка или языка мышления, представляющую определенную идею, понятие. Как справедливо подчеркивал академик Л. В. Щерба, понятие, выражаемое словом (т. е. составляющее смысл этого слова), является правилом лексического употребления данного слова [52]. Никакое слово, взятое вне совокупности всех его лексических употреблений, не выражает адекватно соответствующего понятия. Это означает, что сколь нейтральным в стилистическом отношении ни было бы слово-доминанта, выбранное для обозначения класса условной эквивалентности, оно все-таки имеет свою смысловую ограниченность и специфичность.

Для обозначения дескрипторов (т. е. их репрезентации в плане выражения) правильней всего было бы употреблять не слова естественного языка, а *номинальные определения* понятий, выражаемых этими дескрипторами как классами условной эквивалентности [53]. В нашем контексте номинальным называется такое определение, которое представляет собой ответ на вопрос, что имеют в виду, когда употребляют любое ключевое слово, взятое из данного класса условной эквивалентности. Иными словами, под номинальным определением здесь понимается описательное выражение основного содержания понятия, приписы-

[52] Л. В. Щерба. Преподавание иностранных языков в средней школе. М.—Л., Изд-во Академии пед. наук РСФСР, 1947, стр. 67.

[53] О номинальных определениях см.: Е. К. Войшвилло. Понятие. М., Изд-во МГУ, 1967, стр. 123.

ваемого в качестве смысла всем и каждому ключевому слову, которые включены в определенный класс условной эквивалентности, или конвенция относительно способа употребления ключевых слов, входящих в этот класс. Номинальное определение вводит понятие в обращение. Так, вводя термин «гидролиз», автор курса общей химии поясняет, что под этим термином он будет далее подразумевать химическую реакцию обменного типа, которой сопровождается растворение веществ в воде [54].

Другой возможный способ обозначения дескрипторов — это приведение во взаимно однозначное соответствие им каких-то условных символов, не имеющих собственных значений и никак не связанных с лексической системой естественного языка. В теоретическом плане такой способ также более предпочтителен, чем использование для этой цели слов естественного языка, так как дескрипторы-символы не порождали бы никаких субъективных или даже ложных представлений. Однако использование в качестве имен дескрипторов (классов условной эквивалентности) их номинальных определений сделало бы ИПЯ очень громоздкими, а применение для этой цели символов сильно затруднило бы практическое использование дескрипторных ИПЯ. Поэтому дескрипторы все-таки приходится обозначать словами естественного языка.

При построении словаря дескрипторов группировка ключевых слов в классы условной эквивалентности (т. е. устранение синонимии) производится путем содержательного анализа достаточно обширного списка ключевых слов. Для этого следует предварительно определить условия, необходимые и достаточные для того, чтобы ключевые слова можно было в рамках данного ИПЯ считать эквивалентными друг другу. Очевидно, что эти условия могут быть различными. Они зависят от целого ряда факторов и в первую очередь — от практических требований, предъявляемых к ИПС (например, от необходимой точности и полноты информационного поиска).

Отношение условной эквивалентности между ключевыми словами может быть установлено на основе анализа отношений между понятиями, которые выражаются этими словами. Известно, что между понятиями могут существовать следующие основные виды отношений:

а) отношение равнозначности, которое имеет место тогда, и только тогда, когда объемы понятий полностью совпадают [55];

---

[54] Б. В. Некрасов. Учебник общей химии. М., Госхимиздат, 1957, стр. 138.

[55] Объем понятия — это множество (класс) предметов, каждый из которых обладает всей совокупностью свойств и признаков, отражающих содержание этого понятия.

б) отношение перекрещивания, которое имеет место тогда, и только тогда, когда лишь часть объема одного понятия входит в объем другого понятия и, в свою очередь, лишь часть объема второго понятия входит в объем первого понятия (ср. понятия «книга» и «монография»);

в) отношение подчинения, которое имеет место тогда, и только тогда, когда объем одного понятия составляет лишь часть объема другого понятия (ср. понятия «книга» и «печатная публикация»);

г) отношение внеположенности, которое имеет место тогда, и только тогда, когда объемы понятий полностью исключают друг друга и при этом не исчерпывают области предметов, о которых ведется рассуждение (ср. понятия «книга» и «патентное описание»).

Ключевые слова эквивалентны по смыслу, если между понятиями, которые они выражают, существует отношение равнозначности. Очевидно, что такое отношение возможно лишь между полными (абсолютными) синонимами. Поэтому оно не может быть использовано в качестве основания для группировки в классы условной эквивалентности неполных и тематических синонимов. Для этой цели подходят отношения перекрещивания, подчинения и внеположенности, которые позволяют объединять в классы условной эквивалентности даже такие ключевые слова, в значениях которых имеется хотя бы один общий признак. Именно на этом основании мы при необходимости можем считать условно эквивалентными, например, следующие слова: *амортизатор*, *буфер*, *гибкость*, *демпфер*, *каучук*, *мяч*, *пружина*, *резина*, *рессора*, *упругость*, *эластичность* и т. д.

Общим свойством, позволяющим объединить эти слова в класс условной эквивалентности, является свойство упругости, которое входит в содержание понятия, выражаемого каждым из перечисленных слов. В этой связи необходимо отметить, что в качестве таких объединяющих свойств и признаков могут быть использованы не только те, которые относятся непосредственно к содержанию понятий. Например, может оказаться удобным считать условно эквивалентными ключевые слова, которые всегда встречаются в пределах одной или нескольких смежных фраз текста.

В дескрипторном ИПЯ условно эквивалентными по смыслу можно считать также антонимы — слова с противоположным значением, но в смысловом отношении сходные друг с другом. Антонимы появляются в результате того, что в значении слова какой-то один качественный признак доведен до противоположного. Например: *жесткость — упругость*, *истина — ложь*, *широта — узость* и т. д.

Ранее уже отмечалось, что в качестве дескрипторов целесообразно использовать лишь такие ключевые слова, которые достаточно часто употребляются для координатного индексирования

документов и информационных запросов. Если же ключевое слово встречается очень редко, то его целесообразно включить в более широкий класс условной эквивалентности, с которым данное ключевое слово находится в отношении подчинения. В естественных языках перенос значения, когда вместо вида называют род, а вместо части — целое (или наоборот), называется синекдохой (от *греч.* synekdoché — «соподразумевание»). К такому приему прибегают также при построении рядов тематических синонимов.

Критерием включения ключевого слова в более широкий по смыслу класс условной эквивалентности может служить показатель частоты употребления этого слова в поисковых образах документов. Очевидно, что если ключевое слово, выступающее в роли дескриптора, встретилось за год, например, менее 10 раз, то такое ключевое слово целесообразно включить в более широкий класс условной эквивалентности. Поэтому для окончательного лексикографического редактирования словаря дескрипторов необходимо накопить информацию о частотности употребления отдельных ключевых слов, что возможно лишь в процессе эксплуатационной отладки ИПС.

Опыт практического использования словарей дескрипторов показывает, что для облегчения процедуры координатного индексирования документов и информационных запросов целесообразно каждый дескриптор снабдить обратной отсылкой к тем ключевым словам, которые в данном словаре выступают в качестве синонимов этого дескриптора. При помощи такой отсылки все понятия, выражаемые этими ключевыми словами-синонимами, включаются в понятие, выражаемое соответствующим дескриптором. Ниже приводятся примеры таких обратных отсылок [56]:

НИОБИЙ *вкл.* колумбий [отношение равнозначности]
АМИНОКИСЛОТЫ *вкл.* валин [отношение подчинения]
ЭЛЕКТРОПРОВОДНОСТЬ *вкл.* электрическое сопротивление [отношение перекрещивания]

В этом примере прописными буквами выделены дескрипторы, а строчными даны ключевые слова, выступающие в качестве синонимов этих дескрипторов (*вкл.* — включает). Такой способ выделения дескрипторов мы будем использовать и в дальнейшем.

Итак, при построении словаря дескрипторов сначала составляется список ключевых слов, применяемых для координатного индексирования документов. Затем производится группировка слов, которые условно можно считать синонимами. Из каждой такой группы выбирается одно слово или словосочетание (семантическая доминанта), которым обозначается вся эта группа слов

[56] T. L. Gillum. Compiling a technical thesaurus. «Journal of Chemical Documentation», 1964, v. 4, № 1, стр. 30.

и которое становится дескриптором. Все остальные слова этой группы условно считаются синонимами дескриптора и связываются с ним при помощи отсылочных помет *см.* (смотри) или *употр.* (употребляй) [57]. Выше уже отмечалось, что в роли дескриптора должно выступать наиболее представительное слово из выделенной группы условно эквивалентных слов (т. е. семантическая доминанта) и что группа условно эквивалентных ключевых слов может состоять всего из одного слова или словосочетания, которое тогда само и является дескриптором. Каждый дескриптор при помощи обратных отсылок *вкл.* (включает) связывается со всеми его синонимами. В словаре дескрипторов эта система отсылок имеет следующий вид:

АБСТРАКТНЫЙ *вкл.* отвлеченный
умозрительный
---
отвлеченный *см.* АБСТРАКТНЫЙ
---
умозрительный *см.* АБСТРАКТНЫЙ

## МНОГОЗНАЧНОСТЬ КЛЮЧЕВЫХ СЛОВ И ЕЕ УСТРАНЕНИЕ

Одной из важных особенностей любого естественного языка является то, что слова в нем (в том числе и многие термины) многозначны и могут быть точно поняты лишь в контексте. В поисковых образах документов и в поисковых предписаниях слова-дескрипторы, взятые из естественного языка, выступают без контекста. Поэтому многозначность таких слов совершенно недопустима. Необходимо отметить, что под многозначностью в данном случае понимается способность одного слова иметь несколько разных значений или обозначать несколько разных предметов, а под устранением (снятием) многозначности — формальное закрепление за словом одного, и только одного, совершенно определенного и строго фиксированного значения. Таким образом, в отношении всех слов дескрипторного ИПЯ должны быть выполнены следующие требования, предъявляемые к терминам: каждое слово такого ИПЯ должно иметь одно определенное, специально оговоренное значение и не совпадать в написании ни с каким другим словом этого языка.

В естественных языках причиной многозначности слов (в широком смысле) является равноименность, т. е. употребление одного имени или знака для обозначения не одного, а нескольких разных предметов. Равноименность проявляется в двух языковых явлениях, которые необходимо отличать друг от друга, а именно —

---

[57] Отсылка может быть сделана и при помощи стрелки, например: хлористый натрий → ПОВАРЕННАЯ СОЛЬ.

в полисемии и омонимии. Полисемия — это перенос названия одного предмета на другие предметы, имеющие с этим предметом какие-то общие признаки или свойства. Омонимия — это совпадение (в звучании и написании) разных слов, которые не имеют ничего общего ни по происхождению, ни по функционированию.

*Полисемия и ее устранение.* Как было указано выше, полисемия (или многозначность) [58] представляет собой следствие переноса названия с одного предмета на другой. Такой перенос названия происходит на основе сходства (по форме, цвету, внутренним свойствам и качествам), по смежности (временной, пространственной, логической и т. п.) и по функции. Например:

*звезда* (геометрическая фигура) — *звезда* (небесное тело)
*бюро* (разновидность стола) — *бюро* (учреждение) — *бюро* (руководящий коллегиальный орган)
*крыло* (птицы) — *крыло* (самолета) и т. д.

Необходимо отметить, что существует прямая связь между полисемией слов и относительной частотой их употребления. Эта связь была выявлена Дж. К. Ципфом в результате статистического анализа словарных материалов. Она может быть приблизительно описана формулой:

$$m_r = \sqrt{F_r},$$

где $m_r$ — число разных значений слова и $F_r$ — относительная частота употребления этого слова. По-видимому, данная формула недействительна лишь для нескольких десятков наиболее употребительных слов [59].

Устранение полисемии отобранных для дескрипторного ИПЯ ключевых слов производится лексикографически, т. е. при построении словаря дескрипторов, причем для этого используется такая же методика, как и при построении обычных толковых словарей. В словаре дескрипторов каждое многозначное ключевое слово обязательно снабжается пояснительной пометой. Эта помета разъясняет, в каком именно значении данное ключевое слово используется в дескрипторном ИПЯ. Если многозначное ключевое слово включается в словарный состав ИПЯ не в одном, а в нескольких своих значениях, то для каждого такого значения ключевого слова в ИПЯ строится новое слово, которое в написании не совпадает с исходным словом. Для этого к исходному ключевому слову приписывается специфический индекс. Другой возможный метод устранения полисемии — замена семантических

---

[58] От *греч.* polis — «много» и sema — «знак».

[59] G. K. Zipf. Human behavior and the principle of least effort. An introduction to human ecology. Cambridge, Mass., Addison-Wesley, 1949, стр. 27—31.

вариантов многозначного ключевого слова однозначными словосочетаниями. Примеры устранения полисемии:

1) матрицы (*математические*)
основания (*химические соединения*)
плазма (*крови*)
решетки (*кристаллические*) и т. д.;
2) конденсаторы-А (*электрические*)
конденсаторы-Б (*для пара*)
реакторы-А (*ядерные*)
реакторы-Б (*химические*)
фильтры-А (*газовые*)
фильтры-Б (*оптические*)
фильтры-В (*электрические*) и т. д.;
3) пантографы (*токосъемники*) см. токоснимающие пантографы
пантографы (*чертежные*) см. чертежные пантографы
библиография (*отрасль знания*) см. библиографическая наука
библиография (*перечень публикаций*) см. библиографические списки
ассимиляция (*при обмене веществ*) см. анаболизм
ассимиляция (*синтаксическая*) см. атракция и т. д.

*Омонимия и ее устранение.* Омонимами называются два или несколько слов, которые совпадают между собой как в звучании, так и на письме во всех (или в ряде) им присущих грамматических формах, но имеют совершенно разные значения (от *греч.* homos — «одинаковый» и onyma — «имя»). Значения омонимов осознаются нами как совершенно не связанные между собой и одно из другого невыводимые[60]. Омонимами могут быть лишь слова, принадлежащие к одному и тому же грамматическому классу. Примеры омонимов:

*лук* — оружие и *лук* — растение
*коса* — сельскохозяйственное орудие и *коса* — убор волос
*ключ* — родник и *ключ* — от замка и т. д.

Омонимы могут быть полными и неполными. Полные омонимы совпадают во всех своих грамматических формах, а неполные — лишь в ряде грамматических форм.

От подлинных омонимов необходимо отличать о м о ф о р м ы, а также о м о ф о н ы и о м о г р а ф ы. Омоформы — это слова, которые совпадают между собой в звучании и на письме всегда лишь в отдельных формах (например, *лечу* — 1-е л. ед. ч. от «летать» и «лечить»). Омоформы могут принадлежать как к одному и тому же, так и к разным грамматическим классам. Омофоны — это слова, которые одинаково произносятся, но имеют разные значения и по-разному изображаются на письме (например, *компания — кампания*, *косный — костный*, *пруд — прут* и т. д.). Омографы (или омограммы) — это разные по звучанию и значе-

[60] Л. А. Булаховский. Введение в языкознание, ч. 2. М., Учпедгиз, 1953, стр. 46; Н. М. Шанский. Лексикология современного русского языка. М., «Просвещение», 1964, стр. 41.

нию, но одинаково изображаемые на письме слова и словоформы (например, *за́мок — замо́к*, *му́ка — мука́*, *вы́купать — выкупа́ть* и т. д.).

Словарный состав дескрипторного ИПЯ образуют полнозначные слова, отобранные из соответствующего естественного языка (большинство этих слов — термины). Каждое такое слово используется в ИПЯ в одной, и только в одной, грамматической форме. Например, существительные обычно используются в именительном падеже единственного числа (слова, обозначающие свойства, собственные имена и т. д.) или в именительном падеже множественного числа (все слова, обозначающие перечислимые предметы). Других грамматических форм существительные в ИПЯ не имеют. Очевидно, что эта особенность дескрипторного ИПЯ исключает возможность появления в нем каких-либо омоформ [61].

В дескрипторном ИПЯ не может быть и омофонов, так как этот язык используется лишь в письменном виде. Таким образом, из естественного языка в дескрипторный ИПЯ проникают лишь омонимы и омонимичные омографы.

Снятие омонимии в дескрипторном ИПЯ производится при помощи тех же лексикографических средств, какие применяются для устранения полисемии слов. В дескрипторном словаре все омонимы и омонимичные омографы снабжаются соответствующими пояснительными пометами и отмечаются специфическими индексами (например, цифровыми), устраняющими совпадение омонимов и омонимичных омографов в написании. Например:

графики-1 (*диаграммы*)
графики-2 (*художники*)
нефрит-1 (*воспаление почек*)
нефрит-2 (*минерал*)
органы-1 (*части организма и т. д.*)
органы-2 (*музыкальные инструменты*) и т. д.

Если в дескрипторном словаре каждому слову приводится в однозначное соответствие цифровой, буквенный или буквенно-цифровой шифр (код), то омонимы и омонимичные омографы получают разные шифры.

При построении многоотраслевых словарей дескрипторов может оказаться более удобным другой метод устранения омонимии. При использовании этого метода омониму, непременно снабженному пояснительной пометой, приводится в однозначное соответствие такое словосочетание, которое раскрывает смысл этого омонима, но не совпадает в написании ни с каким другим словом. Это словосочетание включается в словарный состав ИПЯ (т. е. становится дескриптором), а исходный омоним рассматривается

---

[61] Сравни с так называемыми изолирующими языками, например, вьетнамским.

лишь как синоним этого словосочетания в естественном языке. Например:

атласы (*географические*) *см.* географические атласы
атласы (*ткани*) *см.* атласные ткани
руны (*письмена*) *см.* рунические письмена
руны (*эпические песни*) *см.* песни-руны и т. д.

Если омоним имеет синонимы (очень редкий случай), то один из этих синонимов может выполнять функцию устраняющего омонимию слова или словосочетания. Например:

неврит (*воспаление нервов*)
невриты (*отростки нервной клетки*) *см.* аксоны

Необходимо отметить, что омонимия частично устраняется в процессе группировки отобранных для дескрипторного ИПЯ ключевых слов в классы условной эквивалентности. Действующий при этом механизм исключения омонимов из словарного состава ИПЯ почти в точности совпадает с только что описанным. Единственное и малосущественное отличие состоит в том, что синоним или словосочетание, которым в ИПЯ заменяется омоним, в первом случае вводится в этот ИПЯ извне, а во втором случае обычно берется из числа ключевых слов, предварительно отобранных для дескрипторного ИПЯ (это условие не является обязательным).

Из естественного языка в дескрипторный ИПЯ (особенно — узкоотраслевой) чаще всего попадают не все, а лишь одно слово из соответствующей омонимичной группы. Такое слово в дескрипторном словаре обычно не отмечается специфическим индексом и не заменяется синонимичным словосочетанием. Оно лишь снабжается пояснительной пометой, которая раскрывает содержание этого слова и таким образом устраняет его смысловую неясность. Например:

баки (*для жидкостп*)
стопы (*в стихосложении*)
титан (*металл*)
ток (*электрический*)
фрегаты (*птицы*) и т. д.

Очевидно, что в таких случаях мы устраняем не омонимию, а порождаемую ею двусмысленность слов. Подводя итог вышесказанному, следует отметить, что при построении конкретного дескрипторного ИПЯ может быть использован любой из описанных методов устранения омонимии или же все эти методы одновременно.

Выше было указано, что при построении дескрипторных словарей для устранения полисемии и омонимии ключевых слов применяются одни и те же методы. Однако из этого факта было бы неправильно сделать вывод, что в данном случае можно не проводить различия между полисемией и омонимией. Дело в том,

что словоформы или словосочетания, образованные с целью устранения полисемии и омонимии ключевых слов, проявляют разные тенденции при группировке ключевых слов в классы условной эквивалентности. Омонимы — это слова, которые в семантическом плане не имеют ничего общего друг с другом. Поэтому очень мала вероятность того, что словоформы или словосочетания, построенные с целью устранения омонимии ключевых слов, попадут в общий класс условной эквивалентности. Наоборот, словоформы ключевого слова (или заменяющие его словосочетания), построенные с целью снятия его многозначности, представляют собой семантические варианты одного и того же слова. Это обстоятельство значительно повышает вероятность попадания таких словоформ и словосочетаний в общий класс условной эквивалентности.

## ОБЩИЕ ПОНЯТИЯ ТЕЗАУРУСА

Ранее было показано, что для обеспечения возможно большей полноты выдачи при информационном поиске, основанном на использовании дескрипторных ИПЯ, необходимо *избыточное индексирование* документов и/или информационных запросов. Под избыточным индексированием в данном случае понимается дополнение поискового образа документа или поискового предписания дескрипторами, которые по смыслу связаны с основными дескрипторами. При этом более предпочтительным, по-видимому, является избыточное индексирование не документов, а информационных запросов. Основания для такого вывода были изложены ранее.

Для того чтобы можно было производить такое избыточное индексирование, необходимо преобразование алфавитного словаря дескрипторов в нормативный словарь-справочник, в котором были бы эксплицитно выражены важнейшие парадигматические связи между дескрипторами. Такие словари-справочники предназначены для следующих целей:

а) служить руководством для перевода текстов с естественного языка на дескрипторный ИПЯ (путем замены ключевых слов соответствующими дескрипторами);

б) обеспечивать возможность избыточного индексирования документов или информационных запросов;

в) быть пособием, которое бы помогало ищущему информацию находить правильные дескрипторы для выражения его *информационной потребности*.

Особое внимание необходимо обратить на пункт (в) и в связи с этим еще раз подчеркнуть, что одна из основных трудностей (если не самая главная), с которыми приходится иметь дело при информационном поиске, заключается в неадекватности словесного выражения потребителями своей *информационной потребности*. «Узловым моментом в деятельности информационной службы, —

писал Д. Фоскетт, — является не перевод терминов *авторов* [документов] на информационно-поисковый язык. Если бы это было так, насколько просты были бы наши проблемы! Нам нужно было бы лишь составить своего рода таблицу перевода, которую всегда можно держать на современном уровне, включая в нее — в соответствии со структурой указателя — любые новые термины, примененные каким-либо автором. Здесь не возникает никакой трудности; авторский текст — это статическая, законченная единица, которая дает определенный контекст ко всем использованным в ней терминам. Действительная трудность для информационной службы заключается в том, чтобы выражать плохо сформулированную, неполную и изменчивую мысль (thought pattern) потребителя в терминах, обеспечивающих выдачу из хранилища информации, которая способна преобразовать эту мысль в некоторую когерентную систему, соответствующую реальности» [62].

Острота проблемы неадекватного словесного выражения информационных потребностей может быть значительно уменьшена, если создать особый словарь, который бы позволял по смыслу находить слова (дескрипторы), необходимые для выражения этого смысла.

Значение в языке — это отношение между именем и смыслом (при понимании значения в соответствии с так называемым треугольником Огдена и Ричардса). Данное отношение является симметричным (reciprocal and reversible) [63]. Поэтому поиск значения можно начинать либо с имени, либо со смысла. Для поиска значений по именам используются обычные алфавитные словари, а для поиска имен (слов) по их смыслам — так называемые *идеологические* (ideological) или *концептуальные* (conceptual) словари [64]. Основное назначение идеологического словаря (и это подчеркивается даже в названиях многих из таких словарей) — обеспечение выбора наиболее подходящих слов для точного выражения той или иной мысли. В теории информационного поиска идеологические словари получили название *тезаурусов* (от *греч.* thesaurus 1) сокровище, богатство, клад; запас; 2) сокровищница; кладовая, склад; *лат.* сокровищница; кладовая, склад). Однако в лексикографии термин «тезаурус» используется также для обозначения словарей иного типа.

---

[62] D. J. Foskett. Language and classification. «Journal of Documentation», 1965, v. 21, № 4, стр. 275.

[63] S. Ullmann. Semantics. An introduction to the science of meaning. Oxford, Blackwell, 1962, стр. 55—57, 63—64.

[64] Раздел семасиологии, изучающий принципы и закономерности обозначения предметов и выражения понятий лексическими и лексико-фразеологическими средствами языка, называется *ономасиологией* (от *греч.* onoma — имя).

Термин «тезаурус» был впервые применен, по-видимому, флорентийцем Брунетто Латини (1220—1294 гг.) для обозначения систематизированной энциклопедии: свой труд он назвал «Книги о сокровище» (Li Livres dou Trésor). В 1532 г. Ж. Тьерри и Р. Этьенн опубликовали свой «Словарь или Сокровище латинского языка» («Dictionarium, seu Linguae Latinae Thesaurus») — толковый словарь латинского языка, в котором словарные статьи были расположены уже в алфавитном порядке. Наконец, в 1572 г. А. Этьенн выпустил в свет еще один толковый словарь — «Сокровище греческого языка» («Thesaurus linguae Graecae»), составление которого было начато еще его отцом Р. Этьенном. Этот словарь не потерял своего значения до настоящего времени [65].

В работе «Опыт общей теории лексикографии», опубликованной в 1940 г., академик Л. В. Щерба дал следующее определение тезауруса:

«Когда говорят thesaurus, то нынче у нас чаще всего имеют при этом в виду «Thesaurus linguae latinae», предприятие пяти немецких академий, начатое еще в 1900 г. и до сих пор доведенное с пропусками лишь до буквы М. Характерная особенность этого типа словарей состоит в том, что в них приводятся все решительно слова, встретившиеся в данном языке хотя бы один раз (т. е. и все так называемые *hapax*'ы), и что под каждым словом приводятся все решительно цитаты из имеющихся на данном языке текстов (в Thesaurus linguae latinae до 600 г.)» [66].

Далее Л. В. Щерба указывает, что в основе противоположения thesaurus — обычный (толковый или переводный) словарь лежит противоположение «языкового материала» и «языковой системы». Под «языковым материалом» Л. В. Щерба понимает все сказанное, услышанное и понятое на данном языке или, по крайней мере, совокупность всех опубликованных на этом языке текстов, а под «языковой системой» — совокупность правил грамматики и правил словаря, т. е. правил применения слов-понятий к реальной действительности. Очевидно, что словарь такого типа менее всего подходит для информационно-поискового языка с его формализованной лексикой [67].

Другой тип словарей образуют идеологические словари, иногда называемые еще словарями аналогий (dictionnaire analogique), аналитическими, методическими, синтетическими или

---

[65] F. Model. Thesaurus in der Dokumentation. Vielschichtigkeit des Begriffes und historische Entwiklung. «Nachrichten für Dokumentation», 1966, v. 17, № 1/2, стр. 5—6; Е. И. Кацпржак. История книги. М., «Книга» 1964, стр. 94—96.

[66] Л. В. Щерба. Опыт общей теории лексикографии. — В кн.: Л. В. Щерба. Избранные работы по языкознанию и фонетике, т. 1. Л., Изд-во ЛГУ, 1958, стр. 69; *hapax* — слово, форма или употребление, встречающееся только один раз (от *греч.* «единожды»).

[67] Там же, стр. 69.

тематическими словарями. Академик Л. В. Щерба писал, что «словарь каждого языка можно организовать, исходя из фактических форм слов (обычный словарь), располагая их или в алфавитном порядке (алфавитный словарь), или по гнездам (гнездовой словарь), а можно организовать его и исходя от значений, т. е. от понятий, выражаемых фактической формой слов (идеологический словарь). Может показаться, что в последнем случае слова собственно будут разрушены, так как одно и то же слово, имея несколько значений, будет фигурировать в разных местах и что это будет уже не словарь, т. е. не список слов, а список понятий. Однако это неверно. Неправильно думать, что слова имеют по несколько значений: это, в сущности говоря, формальная и даже просто типографская точка зрения. На самом деле мы имеем всегда столько слов, сколько данное фактическое слово имеет значений (так и печаталось, между прочим, в старых словарях: заглавное слово повторялось столько раз, сколько у него было значений). Это вытекает логически из признания единства формы и содержания, и мы должны были бы говорить не о словах просто, а о словах-понятиях. В нашем повседневном употреблении мы скатываемся на формальную точку зрения, придавая слову *слово* значение «фонетического слова». Таким образом, точнее всего было бы говорить, что обычный словарь является списком «фонетических слов» с их значениями, а идеологический словарь является списком слов-понятий с их синонимами» [68].

Для создания идеологического словаря прежде всего необходимо точно описать значения всех употребляемых в данном языке слов и затем составить полный список слов-понятий этого языка. Создание идеологического словаря завершается классификацией слов-понятий, которая должна раскрывать объективные взаимосвязи между этими словами-понятиями. Теоретические основы и принципы составления идеологических словарей, по-видимому, наиболее полно изложены в работе Р. Халлига и В. Вартбурга «Система понятий как основа лексикографии» [69].

Из идеологических словарей античного периода наиболее известным является «Ономастикон» Поллукса, основанный на более ранних словарях Аристофана Византийского, Памфилия и Дидима Халкентера. В этом идеологическом словаре (тезаурусе) слова греческого языка (включая все синонимы) сгруппированы в десять следующих классов: Боги, Человек (части тела), Родство, Наука и искусство, Охота, Еда, Ремесла, Закон, Городская организация, Утварь. В наше время широкую известность получил «Thesaurus of English Words and Phrases», который был составлен в 1852 г. англичанином П. М. Роджетом «для облегчения

[68] Там же, стр. 77—78.

[69] R. Hallig und W. von Wartburg. Begriffssystem als Grundlage für Lexikographie. Berlin, 1952.

выражения понятий (ideas) и для помощи при написании сочинений»[70]. С тех пор этот словарь переиздавался сыном и внуком автора не менее 90 раз.

Необходимо отметить, что на тезаурус Роджета оказал влияние тезаурус санскрита, опубликованный в 1808 г. В этом тезаурусе слова сгруппированы в следующие классы: *Небеса*, *Погода*, *Время*, *Добродетели*, *Пороки*, *Знания*, *Чувства*, *Речь*, *Музыка*, *Ад*, *Вода*, *Земля*, *Растения*, *Животные*, *Человек*, *Семья*, *Медицина*, *Одежда*, *Раса*, *Племя*, *Правительство*, *Армия*, *Занятия*, *Товары*, *Утварь*.

В свою очередь, эти классы делятся на соответствующие подклассы. Например, класс *Ад* имеет подразделения *Ад как такового*, *Души умерших*, *Страдание*, *Боль*[71].

Тезаурус Роджета в известном смысле является обращенным вариантом толкового словаря. В толковом словаре слова располагаются в алфавитном порядке и значения этих слов раскрываются примерами их конкретных употреблений. Наоборот, тезаурус Роджета, как отмечал сам его составитель в предисловии к первому изданию этого словаря, предназначен для того, чтобы «к каждой приводимой в нем идее можно было подобрать слово или слова, с помощью которых эту идею можно выразить наиболее правильно и точно. С этой целью слова и выражения английского языка систематизированы здесь не по их произношению и написанию, а строго в соответствии с их значением»[72].

В 1956 г. тезаурус Роджета был использован в работах по автоматическому переводу Кембриджской группой по исследованию языка. Члены этой группы А. Ф. Паркер-Роудс и С. Уордли дают следующее определение понятия «тезаурус»: «На вопрос, что такое тезаурус, проще всего было бы ответить, показав несколько существующих словарей — тезаурусов, например, хорошо известный словарь Роджета, и сказать, что тезаурус — это любая система классификации слов языка, относящаяся к такому же типу и имеющая такую же форму, что и этот словарь. Для тех, кто не знаком с подобными словарями, можно сказать, что имеется в виду обращенный словарь. Если в обычном словаре по слову находят его значение, обычно определяемое рядом примеров его употребления, то в тезаурусе по значению, которое записывается некоторым способом, находят слово или несколько слов, выражающих это искомое значение. По существу, тезау-

[70] P. M. Roget. Thesaurus of English Words and Phrases classified and arranged so as to facilitate the expression of ideas and to assist in literary composition. London, 1852.

[71] Amera Sinha. Cosha, or Dictionary of the Sanskrit language. Serampoor, 1808. (См. кн.: K. Sparck Jones. Synonymy and semantic classification. (Ph. D. Thesis, University of Cambridge, 1964). Rep. NOML 170. Cambridge, Cambridge Language Research Unit., 1964, app. 1, стр. 2.

[72] P. M. Roget. Thesaurus of English words and phrases. Abridged ed. Harmonsworth, Penguin Books, 1960, стр. V.

рус, так же как и обычный словарь, есть способ определения значения слов в терминах их возможных употреблений. Но принципы размещения материала в этих двух типах лексикографических работ противоположны»[73].

Тезаурус Роджета представляет собой собрание слов английского языка, сгруппированных в 6 крупных тематических классов (категорий). Эти классы подразделяются на 24 подкласса, которые, в свою очередь, делятся на 86 подподклассов, и т. д. Ниже перечислены названия всех 6 категорий и 24 классов тезауруса Роджета, а также названия последующих подразделений, на которые делится класс 1.4. *Упорядоченность*.

1. Абстрактные отношения
1.1. Бытие
1.2. Отношение
1.3. Количество
1.4. *Упорядоченность* → 1.4.1. Упорядоченность вообще
1.5. Число → 1.4.2. Консекутивный порядок
1.6. Время → 1.4.3. Совокупный порядок
1.7. Изменение
1.8. Причинность

2. Пространство
2.1. Пространство вообще → 1.4.4. *Дистрибутивный порядок*
2.2. Измерения
2.3. Форма
2.4. Движение

75. Класс
76. Включение 77. Исключение
78. Всеобщность 79. Особенность

3. Материя
3.1. Материя вообще → 1.4.5. Категорический порядок
3.2. Неорганическая материя
3.3. Органическая материя

4. Разум
4.1. Образование понятий
4.2. Передача идей

5. Воля
5.1. Индивидуальная воля
5.2. Общественная воля

6. Чувственные и моральные силы
6.1. Чувства вообще
6.2. Индивидуальные чувства
6.3. Общественные чувства
6.4. Моральные чувства
6.5. Религиозные чувства

Всего в тезаурусе Роджета на низшем уровне иерархии насчитывается 1000 тематических групп[74]. В каждой тематической

[73] А. Ф. Паркер-Роудс и С. Уордли. Применение тезаурусного метода при машинном переводе с помощью существующей машинной техники. «Математическая лингвистика. Сборник переводов». М., «Мир», 1964, стр. 215.

[74] В классификации слов-понятий, которую на материале французского языка разработал Ш. Балли, выделено 10 категорий, 41 класс и т. д.; в конеч-

группе приводятся близкие по смыслу слова английского языка, которые могут быть использованы для выражения понятия, обозначенного «заглавным» словом. Вследствие полисемии и омонимии слово может одновременно входить в несколько тематических групп. Например, слово «knot» входит в следующие тематические группы:

45 — связь
59 — запутанность
72 — группа
219 — пересечение
249 — округлять(ся)
321 — плотность
704 — трудность и т. д.[75]

Тезаурус Роджета снабжен указателем, в котором все включенные в него слова расположены в алфавитном порядке и каждому слову приписаны номера тематических групп, в которые входит это слово. Слова, обозначающие тематические группы, выделены в этом указателе полужирным шрифтом.

После появления тезауруса Роджета аналогичные словари начали составляться и на других языках. Ниже перечислены основные словари такого типа.

R. Bossière. Dictionnaire analogique de la langue française. Répertoire complet des mots par les idées, des idées par les mots. Paris, 1862.
F. Dornseiff. Der deutsche Wortschatz nach Sachgruppen. Berlin, 1933.
Ch. Marquet. Dictionnaire analogique. Répertoire moderne des mots par les idées, des idées par les mots. Paris, 1936.
J. Casares. Diccionario ideológico de la lengua española. Desde la idea a la palabra; desde la palabra a la idea. Barcelona, 1951.
P. Robert. Dictionnaire alphabétique et analogique de la langue française. Les mots et les associations d'idées. v. 1—4. Paris, 1953—1959.
T. Robertson. Dictionnaire idéologique. Recueil des mots, des phrases, des idiotismes et des proverbes de la langue française, classés selon l'ordre des idées. Paris, 1959.
P. Rouaix. Dictionnaire des idées suggerées par les mots. 15 ed. Paris, 1926 (первое издание вышло в 1911 г.).
D. Sanders. Deutscher Sprachschatz geordnet nach Begriffen zur leichten Auffindung und Auswahl des passenden Ausdrucks. Hamburg, 1873—1877.
Schlessing. Der passende Ausdruck. Berlin, 1881. (В 1927 г. этот словарь вышел в переработанном виде: Schlessing und H. Wehrle. Deutscher Wartschatz. Ein Hilfs- und Nachschlagebuch der sinnverwandten Wörter und Ausdrücke der deutschen Sprache).

В словарях Р. Босьера и Ш. Марке каждая группа слов, связанных по смыслу (par analogy), обозначена определенным

ном итоге в этой классификации насчитывается 297 видовых подразделений. См. Ch. Bally. Traité de stylistique française. 3-me ed., v. 1—2. Geneva-Paris, 1951.

[75] P. M. Roget. Thesaurus of English words and phrases. Abridged ed. Harmonsworth, Penguin Books, 1960.

или слова, входящие в типичные сочетания с заглавным словом. Ключ — это список всех слов в алфавитном порядке с указанием номеров рубрик, в к[ото]рые слова входят». ‹. . . .› «Бывает целесообразно рассматривать и более общее понятие т[езауруса] как списка объектов, предикатов и связей между ними в определ[енной] сематич[еской] системе» [82].

3. «Технический тезаурус — это переплетенный том, в котором физически воплощены методы лексикографического контроля, применяемого в данной информационной системе. Для решения проблем точки зрения, семантики и родовидовых связей в тезаурусе используются пометы «связанные термины» (related terms), «смотри» и «приписать родовой термин» (generic posting) ‹. . .› Технический тезаурус — это такое средство (device), которое позволяет производить либо избыточное индексирование, либо избыточный информационный поиск. Избыточность можно применять продуманно и избирательно, в зависимости от ценности хранимого материала или от важности полученного запроса. Технический тезаурус позволяет как индексатору, так и проводящему информационный поиск применять такую степень избыточности, которая необходима в данной конкретной ситуации» [83].

4. «Тезаурус можно определить как список ведущих (lead) терминов, под каждым из которых приводятся связанные с ним термины, к которым предметизатор обращается после первоначального индексирования. Он изучает список ассоциативных терминов, приведенных под каждым из терминов, которыми он заиндексировал документ, и из таких списков выбирает дополнительные термины с тем, чтобы обеспечить более адекватное описание, даваемое информационному содержанию документа. Тезаурус обращает его внимание на термины, которые дадут ему возможность описать информацию с других точек зрения» [84].

«Технический тезаурус функционирует как список для напоминания слов (word-reminder list). Он дает возможность индексаторам: (1) описывать индексируемую информацию столько числом терминов, сколько следует или необходимо для выражения разных точек зрения, с которых может рассматриваться содержащаяся в документе информация; (2) описывать информацию, когда это необходимо, в синонимичных и полусинонимичных терминах

---

[82] Ю. А. Шрейдер. Тезаурус. «Автоматизация производства и промышленная электроника», т. 3. М., «Сов. энциклопедия», 1964, стр. 473. См. также: Ю. А. Шрейдер. О количественных характеристиках семантической информации. Сб.: «Научно-техн. информация», 1963, № 10, стр. 33—38.

[83] B. E. Holm and L. E. Rasmussen. Development of a technical thesaurus. «American Documentation», 1961, v. 12, № 3, стр. 186.

[84] J. C. Costello, Jr. Uniterm indexing principles, problems and solutions. «American Documentation», 1961, v. 12, № 1, стр. 23.

и (3) относить узкие понятия к более широким, находящимся на более высоких уровнях иерархии» [85].

5. «Тезаурус — это собрание (collection) таких слов как синонимы и антонимы, которые систематизированы по «понятиям» — сравнительно широким смысловым гнездам (clusters), а не по алфавиту или другим характеристикам. В документации термин «тезаурус» иногда используется для обозначения списка нормативных (authorized) предметных заголовков, составленного таким образом, что сходные [по смыслу] слова собраны в группы и снабжены ссылками «смотри» и «смотри также» [86].

6. «Термин — „тезаурус" в теории информационного поиска может быть определен как «собрание нормативных предметных заголовков или «дескрипторов». Эти дескрипторы систематизированы по понятийным группам и полям, сопровождаются алфавитным указателем и содержат необходимые пояснительные полеты и перекрестные ссылки» [87].

7. «Тезаурус — это словарь (lexicon), особенно такой, в котором слова сгруппированы по смыслу (by ideas); группировка или классификация синонимов и полусинонимов; множество терминологических классов эквивалентности» [88].

8. «Тезаурус как вспомогательное средство документации представляет собой собрание слов естественного языка (общего и/или специального) с обозначением смысловых связей между этими словами. Тезаурус может быть построен по систематическому принципу или — с помощью соответствующих ссылок — по какому-либо формальному (например, алфавитному) принципу. В отличие от других классификационных систем в тезаурусе все слова одной понятийной области по возможности группируются вместе, причем допустимо применение ссылок на принятые классификационные символы (слова или шифры)» (Определение понятия «тезаурус» Немецкого общества документации, ФРГ).

9. «Назначение информационно-поискового тезауруса (information retrieval thesaurus) заключается в следующем: (1) дать лицу, производящему индексирование документов, . . . возмож-

---

[85] J. C. Costello, Jr. Training manual and workbook for use in abstracting and coordinate indexing. Training course. Columbus, Battelle Memorial Institute, 1964, стр. 101.

[86] R. S. Taylor. Glossary of terms frequently used in scientific documentation. New York, American Institute of Physics, 1962, стр. 15; D. L. Thompson. Glossary of STINFO terminology. Dayton, Ohio, Office of Aerospace Research (AFOSR 5266), 1963, стр. 129.

[87] C. F. Balz. The need for a thesaurus in automated information retieval. Owego, N. Y., IBM Space Guidance Center, 1962, стр. 1.

[88] Reference manual. Index organization for information retrieval. White Plains. N. Y., IBM Technical Publications Department, 1961, стр. 58; W. F. Williams. Principles of automated information retrieval. Elmhurst, Ill., The Business Press, 1965, стр. 425.

ность индексировать (т. е. описывать) содержащуюся в этих документах информацию более полно, на разных уровнях обобщения (generality) и в техническом смысле — с многих точек зрения и (2) позволять лицу, ведущему поиск информации, формулировать поисковое предписание в соответствии с объемом и степенью его непосредственных интересов, т. е. так, чтобы поисковое предписание содержало все термины словарного состава информационно-поискового языка, которые обладают нужным значением и специфичностью» [89].

10. «Тезаурус — это словарь абсолютных или условных эквивалентностей между словами или выражениями естественного языка и «дескрипторами» информационно-поискового языка (langage documentaire)» [90].

11. «Тезаурус — это списки соответствий между терминами или группами терминов информационно-поискового языка. . . и словами или выражениями естественного языка» [91]. Эти термины или группы терминов информационно-поискового языка могут принадлежать либо к множеству специальных терминов (ключевых слов или дескрипторов), которые не упорядочены по смысловым признакам, но могут быть упорядочены по алфавиту (vocabulaire), либо к алфавитному списку специальных терминов, которые снабжены определениями и используются для целей индексирования (glossaire documentaire), либо к классификации. Под классификацией в данном случае понимается любое множество терминов, систематизированное по их смысловым признакам [92].

12. «Тезаурус — это список, в котором перечислены понятия, выраженные на естественном языке (обычные и/или научно-технические), и обозначены отношения между ними» [93].

13. «Тезаурус в самом широком смысле — это просто классификация слов по понятиям, темам или предметам. . .» [94].

Анализ приведенных выше определений позволяет сделать выводы о назначении и общей структуре тезауруса, предназначенного для целей информационного поиска. Прежде всего необходимо отметить, что такой тезаурус не является лишь идеологиче-

---

[89] E. Wall. Information retrieval thesauri. New York, Engineers Joint Council, 1962, стр. 1.

[90] R. C. Cros, J.-C. Gardin, F. Lévy. L'automatisation des recherches documentaires. Un modèle général le Syntol. Paris, Gauthier-Villars, 1964, стр. 254.

[91] N. Gardin. Traitement automatiques des donnees non numeriques. Projet de lexique documentaire. Revision № 1, Chapitres 1 et 2 (Annexes). Marseille, Groupe d'Etude sur l'Information Scientifique, Section d'Automatique Documentaire (C.N.R.S.), 1965, стр. 40.

[92] Там же, стр. 39—40.

[93] FID/ISO Project: Classification Terminology. Document DH/6F (French). Copenhagen, FID/CR Committee on Classification Research, 1967.

[94] K. Sparck Jones. Synonymy and semantic classification. (Ph. D. Thesis, Univ. of Cambridge). ML 170. Cambridge, Cambridge Language Research Unit., 1964, App. 1, стр. 1.

ским словарем дескрипторов, так как он должен выполнять также функцию двуязычного словаря (естественный язык — дескрипторный ИПЯ), обеспечивающего перевод текстов поисковых образов и информационных запросов с естественного языка на дескрипторный ИПЯ (и наоборот). Это означает, что информационно-поисковый тезаурус должен содержать не только все дескрипторы, но и все ключевые слова, которые в рамках данной ИПС считаются синонимами дескрипторов.

Кроме того, информационно-поисковый тезаурус, как уже отмечалось в начале данного раздела, является нормативным пособием, предназначенным для предметизаторов, индексаторов и потребителей информации. Поэтому такой тезаурус наиболее целесообразно выполнять в виде одно- или двухтомного справочника, содержащего две основные части:

А. Общий алфавитный список дескрипторов и ключевых слов, снабженных соответствующими пометами и ссылками. В этом списке ключевые слова могут быть определенным образом отмечены, чтобы их легче было отличать от дескрипторов. Под каждым дескриптором должны быть приведены в упорядоченном виде другие дескрипторы, которые по смыслу связаны с «заглавным» дескриптором [95]. Желательно, чтобы типы этих смысловых связей были соответствующим образом обозначены. В словарной статье целесообразно также дать список всех ключевых слов, которые в рамках данной ИПС считаются синонимами «заглавного» дескриптора.

Б. Упорядоченное множество тематических классов (полей) и/или групп дескрипторов. Каждый тематический класс (поле) представляет собой подмножество наименований тематических групп, а каждая тематическая группа — подмножество дескрипторов, которое обозначено, как правило, одним или несколькими дескрипторами, выбранными из данного подмножества дескрипторов. В качестве наименований тематических групп и особенно классов (полей) могут использоваться слова и словосочетания естественного языка, не являющиеся дескрипторами. Внутри каждого тематического класса (поля) или группы наименования групп и дескрипторы также должны быть упорядочены (например, по алфавиту и/или по какому-либо иному признаку). Смысловые связи между дескрипторами могут быть выражены лексикографически, в виде таблиц (списков), аналитически или графически.

Необходимо отметить, что части (Б), как будет показано далее, имеют не все тезаурусы. Отсутствие в информационно-поисковом тезаурусе части (Б) с упорядоченной системой тематических классов (полей) и/или групп дескрипторов значительно затруд-

[95] Возможна замена всех этих дескрипторов индексом тематической группы, в которую входит «заглавный» дескриптор.

няет использование такого тезауруса в функции идеологического словаря, т. е. для более адекватного выражения в терминах дескрипторного ИПЯ фактической информационной потребности того, кто проводит поиск информации.

Отсутствие части (Б) делает информационно-поисковый тезаурус весьма похожим на упорядоченный по алфавиту нормативный список предметных заголовков. Поэтому информационно-поисковые тезаурусы иногда называют «нормативными словниками предметных заголовков» (authorized subject-heading list). Но это сходство чисто внешнее, так как *построение дескрипторов как классов условной эквивалентности основано на совершенно иных принципах, чем формулирование заголовков и подзаголовков для словников алфавитно-предметных классификаций.*

В алфавитно-предметных классификациях, которые при информационном поиске не допускают логического умножения классов, обозначенных предметными заголовками, каждый такой заголовок должен обозначать *исчерпывающий класс* предметов и явлений, обладающих той или иной *совокупностью каких-то общих признаков.* Такие предметные заголовки составляют очевидное большинство в любой алфавитно-предметной классификации. Что же касается дескрипторов, то они должны обозначать, как правило, классы предметов и явлений, обладающие каким-то *одним общим признаком.* Обычно такие классы по своему объему значительно шире тех, которые обозначаются предметными заголовками. Однако путем логического умножения этих широких классов можно строить новые классы, имеющие почти любой малый объем.

Например, для того чтобы по алфавитно-предметной классификации заиндексировать документ о производстве и экспорте вычислительных и пишущих машин из Великобритании, Италии и Франции, необходимы такие обобщенные предметные заголовки:

Производство — вычислительные машины — Великобритания
Производство — вычислительные машины — Италия
Производство — вычислительные машины — Франция
Производство — пишущие машины — Великобритания
Производство — пишущие машины — Италия
Производство — пишущие машины — Франция
Экспорт — вычислительные машины — Великобритания
Экспорт — вычислительные машины — Италия
Экспорт — вычислительные машины — Франция
Экспорт — пишущие машины — Великобритания
Экспорт — пишущие машины — Италия
Экспорт — пишущие машины — Франция

Под обобщенным предметным заголовком в данном случае понимается совокупность предметного заголовка, предметного подзаголовка (одного или нескольких) и предметной записи. В связи с этим необходимо отметить, что окончательное формирование предметных классов происходит под сильным влиянием

смыслового содержания индексируемых документов. Вследствие этого система предметных классов (а следовательно, и система обобщенных предметных заголовков) приспосабливается к содержанию массива заиндексированных документов, что снижает ее разрешающую способность относительно «неожиданных» информационных запросов.

Для координатного индексирования упомянутого выше документа достаточно иметь лишь следующие дескрипторы:

ВЫЧИСЛИТЕЛЬНЫЕ МАШИНЫ, ПИШУЩИЕ МАШИНЫ, ПРОИЗВОДСТВО, ЭКСПОРТ, ВЕЛИКОБРИТАНИЯ, ИТАЛИЯ, ФРАНЦИЯ.

Полезно напомнить, что в словнике алфавитно-предметной классификации, которая до 1960 г. использовалась в Управлении технической информации вооруженных сил США (ныне — Военный центр документации — Defense Documentation Center — DDC), насчитывалось около 70 тыс. предметных заголовков и подзаголовков, а составленный на основе этого словника тезаурус (первое издание) содержал менее 7 тыс. дескрипторов [96].

Все сказанное выше позволяет сделать вывод, что вследствие неэквивалентности понятий «предметный заголовок» и «дескриптор» не эквивалентны также и понятия «нормативный список предметных заголовков» и «тезаурус», даже если в данном случае имеется в виду упрощенный тезаурус, в котором отсутствует раздел с тематическими классами и/или группами дескрипторов. Это означает, что *нормативные списки предметных заголовков не могут заменять тезаурусы*, хотя их использование при построении соответствующих тезаурусов следует признать весьма полезным.

Этот вывод целиком распространяется и на иерархические библиотечно-библиографические классификации (в дальнейшем мы будем называть их просто иерархическими классификациями). В связи с этим необходимо уточнить, что в данном случае речь идет не об абстрактной возможности использования классов (или имен классов) алфавитно-предметной или иерархической классификации в качестве дескрипторов. Очевидно, что для такого их использования нет никаких формальных препятствий. Но возникает вопрос, целесообразно ли это. На такой вопрос может быть дан лишь отрицательный ответ.

Обоснование такого ответа представляется полезным начать с рассмотрения причин, поставивших на обсуждение сформулированный выше вопрос. Как показывают теория и практика, алфавитно-предметные и иерархические классификации недостаточно эффективны при поиске документов по любому, заранее не пред-

[96] J. J. Jaster, B. R. Murray and M. Taube. The state of the art of coordinate indexing (Preliminary Edition). Washington, Documentation, 1962, стр. 40—41.

усмотренному сочетанию характеристик. Эта особенность алфавитно-предметных и иерархических классификаций была подробно рассмотрена в главе 5. Лучшие результаты дает использование для этой цели фасетных классификаций. Дескрипторные ИПЯ значительно превосходят алфавитно-предметные, иерархические и фасетные классификации при поиске документов по информационным запросам такого типа. Но создание дескрипторных языков — дело весьма нелегкое. Между тем алфавитно-предметные и иерархические классификации разрабатываются и используются уже давно. Такие классификации имеются, по-видимому, почти по любой отрасли науки и техники. Координатное индексирование, лежащее в основе дескрипторных ИПЯ, — это тоже разновидность классификации документов. Поэтому возникает естественный вопрос, нельзя ли использовать подразделения уже составленных алфавитно-предметных, иерархических и фасетных классификаций в качестве дескрипторов.

Особенно соблазнительным представляется использование в «дескрипторном режиме» подразделений УДК, так как эта классификация охватывает все отрасли науки и техники, имеет международное признание и является в целом наиболее совершенной из всех существующих систем библиотечно-библиографических классификаций. Кроме того, для постоянного совершенствования УДК в Международной федерации по документации уже давно создан уникальный, хотя и медленно действующий механизм внесения изменений и дополнений в таблицы этой классификации. По УДК уже заиндексированы обширные собрания научно-технических документов и накоплен большой опыт практического использования этой классификации. Наконец, по УДК индексируются публикации, помещаемые в сотнях зарубежных периодических изданий, а также все публикации по науке и технике, выходящие в нашей стране (в соответствии с Постановлением Совета Министров СССР от 11 мая 1962 г. «О мерах по улучшению организации научно-технической информации в стране» УДК введена в качестве обязательной для классифицирования научно-технической литературы во всех издательствах, научно-технических библиотеках и органах информации страны).

Однако при всей заманчивости такой перспективы нельзя забывать о том, что алфавитно-предметные и иерархические классификации создавались для вполне конкретных целей и способов использования, которые во многом не совпадают с назначением и способами использования дескрипторных ИПЯ. Едва ли целесообразно «наливать вино молодое в мехи старые». Ведь использование подразделений алфавитно-предметных или иерархических классификаций в качестве дескрипторов означает, что в лучшем случае мы получаем плохой суррогат дескрипторного ИПЯ, но лишаемся таких важных достоинств алфавитно-предметных и иерархических классификаций, как простота создания и исполь-

зования основанных на этих классификациях поисковых систем — алфавитно-предметных и систематических каталогов (указателей), а также возможность корректировки хода информационного поиска в этих системах. Если прибегнуть к аналогии, то предложения использовать алфавитно-предметные и иерархические классификации в «дескрипторном режиме» в какой-то мере подобны предложению резать листовое железо садовыми ножницами, причем сделанному только на том основании, что у нас уже имеются такие ножницы, тогда как специальные ножницы для резки металла нужно еще приобрести.

Как уже указывалось ранее, одно из основных соображений, которыми руководствуются при построении дескрипторов как классов условной эквивалентности, состоит в том, что в процессе информационного поиска над этими классами выполняются операции логического умножения, сложения и отрицания. Наоборот, при создании алфавитно-предметной или иерархической классификации важнейшей задачей является то, чтобы иметь в ней исчерпывающие классы для любых предвидимых сочетаний характеристик и таким образом исключить необходимость каких-либо логических операций над этими классами в процессе информационного поиска. Поэтому a priori можно утверждать, что общее число подразделений любой алфавитно-предметной или иерархической классификации должно быть во много раз больше числа дескрипторов в эквивалентном ей по предметному охвату дескрипторном ИПЯ. Вместе с тем столь же очевидно, что «словарный состав» любой такой классификации (т. е. совокупность имен ее классов и подклассов) будет всегда беднее словарного состава соответствующего ей дескрипторного ИПЯ. В отношении иерархических классификаций необходимо также учитывать трудность построения классов для новых и особенно многоплановых понятий. Из сказанного следует, что алфавитно-предметные и иерархические классификации, используемые в «дескрипторном режиме» по семантической силе, будут значительно уступать подлинно дескрипторным ИПЯ.

Иначе обстоит дело с использованием в качестве дескрипторов подразделений фасетных классификаций. Принципы построения фасетных классификаций таковы, что позволяют использовать таблицы этих классификаций как тезаурусы, а содержащиеся в фасетах термины (фокусы) — как дескрипторы.

В фасетных классификациях термины систематизированы (сгруппированы в фасеты) и снабжены специфическими шифрами. Это в значительной степени устраняет лексическую многозначность терминов, включенных в таблицы фасетной классификации. Группировка терминов по категориям (фасетам) позволяет собрать в одном месте все термины, выражающие соподчиненные (видовые) понятия. При алфавитном расположении в едином списке эти термины были бы рассеяны по всему списку. Кроме того, термины,

образующие тот или иной фасет, выполняют функцию контекста, который как бы уточняет в этом фасете значение любого термина и определяет его синтаксическую роль. Данное обстоятельство делает ненужным применение специальных указателей роли [97], о которых будет идти речь далее. В приведенном на стр. 340 примере шифр *Lm* показывает, что термин «удар» принадлежит к фасету *L* (напряжения и нагрузки) и поэтому означает не удар вообще, а удар как приложенную нагрузку. Наконец, систематическая расстановка карточек в каталоге, организованном по инвертированной схеме, позволяет в процессе поиска расширять или сужать информационный запрос в зависимости от количества документов, выдаваемых поисковой системой. При таком использовании таблиц фасетной классификации мы имеем дело фактически уже не с классификацией, а с ИПЯ дескрипторного типа.

Вывод о возможности использования классов (или имен классов) фасетных классификаций в качестве дескрипторов распространяется и на УДК, которая является классификацией полуфасетного типа. Однако иерархический принцип построения общепредметного фасета УДК (т. е. десяти ее главных классов) значительно ограничивает такую возможность, так как в отношении этого фасета УДК в основном справедливы все те критические замечания, которые были сделаны выше в отношении алфавитно-предметных и иерархических классификаций. К такому существенному недостатку общепредметного фасета УДК, как чрезмерно большой объем «словарного состава» (при недостаточной его полноте) следует добавить еще неустранимую синонимию индексов этого фасета, а также неоднозначность правил построения.

Для того чтобы получить общее представление о соотношении объемов «словарного состава» УДК и словаря эквивалентного ему по предметному охвату дескрипторного ИПЯ, достаточно привести следующие цифры. В полном издании УДК класс 6 «Прикладные науки. Медицина. Техника» содержит 64 тыс. подразделений (а следовательно, и их наименований), тогда как в «Тезаурусе технических терминов», который был составлен и опубликован в США (1964 г.), насчитывается всего 7750 дескрипторов [98]. При этом можно с уверенностью утверждать, что сло-

---

[97] Термины «указатели роли» и «инвертированная схема организации каталога» определены соответственно на стр. 507 и 524.

[98] M. Rigby. Mechanization of the UDC. Final report on pilot project to further explore possibilities for mechanization of Universal Decimal Classification (UDC) schedules. Washington, American Meteorological Society, Meteorological and Geoastrophysical Abstracts, 1963, стр. 4; D. J. Campbell. Thesaurus of engineering terms: a list of engineering terms and their relationships for use in vocabulary control, in indexing and retrieving engineering information. New York, Engineers Joint Council, 1964. [Рецензия]. «Journal of Documentation», 1965, v. 21, № 2, стр. 137.

варный состав этого дескрипторного ИПЯ значительно богаче «словарного состава» класса 6 УДК. Определенная часть подразделений класса 6 УДК относится к медицине, терминология которой не отражена в «Тезаурусе технических терминов». Но этот тезаурус охватывает еще значительную часть класса 5 УДК «Математика. Естественные науки», в котором насчитывается около 35 тыс. подразделений.

Итак, УДК может применяться в «дескрипторном режиме», так как она имеет тщательно систематизированный и достаточно исчерпывающий словарь терминов; эти термины или, по крайней мере, многие из них выражают достаточно простые понятия, что позволяет широко применять их посткоординацию — этот modus operandi ИПС дескрипторного типа; наконец, в индексах УДК, которые можно использовать непосредственно в качестве кодов терминов, эксплицитно выражены родо-видовые связи между этими терминами. Однако вновь возникает все тот же вопрос, *следует ли применять УДК в «дескрипторном режиме»*. На такой вопрос известный английский специалист по УДК Дж. Миллз дает следующий ответ: «Вообще говоря, организованный по УДК указатель не требует никакой механизации [выполнения] его функций индексирования и поиска, возможно, за исключением случаев применения УДК в некоторых областях, например для поиска химических соединений, когда преимущества посткоординации перед предкоординированными классами не могут вызывать никаких сомнений» [99].

В начале 1967 г. американские специалисты Р. Фримен и П. Атертон провели интересный эксперимент по использованию УДК в «дескрипторном режиме». Для эксперимента был использован массив примерно из 250 документов по океанографии, заиндексированных по УДК. Поисковая система была реализована на ЭЦВМ типа ИБМ-1401. Результаты этого эксперимента позволили Р. Фримену и П. Атертон сделать следующие общие выводы:

«Первое и самое очевидное: нет более никаких сомнений, что УДК может быть использована в механизированной системе. Но нужно найти ответ на более трудный вопрос о том, *следует* ли ее использовать [в таком режиме]» [100]. И далее: «Наш второй вывод . . . . заключается в следующем: УДК в том виде, в каком она существует в настоящее время, вероятно, не может функциони-

---

[99] J. Mills. The Universal Decimal Classification. New Brunswick, N. Y., The Rutgers University Press, 1964, стр. 14. (Rutgers series on systems for the intellectual organization of information, v. 1).

[100] R. R. Freeman and P. Atherton. File organization and search strategy using UDC in mechanized reference retrieval systems. (Preprint of a paper for presentation at the F.I.D./I.F.I.P. Conference on mechanized information storage, retrieval, and dissemination. Rome, June, 15, 1967), [New York], American Institute of Physics, 1967, стр. 17.

ровать в механизированной системе столь же эффективно, как и язык индексирования, специально предназначенный для машинной обработки; но не существует никаких препятствий для успешного использования УДК либо в режиме циклической обработки (batch processing), либо в режиме взаимодействия между человеком и машиной (interactive mode)» [101].

В заключение данного раздела полезно подвести некоторые итоги. Прежде всего необходимо указать, что информационно-поисковые тезаурусы могут быть трех типов: лингвистические, статистические и нормативные. *Лингвистическим тезаурусом* называется некоторое множество слов естественного языка, отобранных путем содержательного анализа текстов и систематизированных в соответствии с предварительно выбранной или разработанной системой классификации. Термин «лингвистический тезаурус» является синонимом термина «идеологический словарь». *Статистическим тезаурусом* называется упорядоченное множество ключевых слов, отобранных путем статистического анализа текстов по какой-либо тематике. *Нормативным тезаурусом* называется словарь-справочник, который предназначен для оказания помощи потребителю в более точном выражении (в терминах дескрипторного ИПЯ) содержания нужной ему информации и для обеспечения возможности избыточного индексирования (также в терминах ИПЯ) документов и информационных запросов. Такой словарь-справочник должен содержать все лексические единицы ИПЯ — дескрипторы (вместе с ключевыми словами, которые в пределах данной ИПС считаются синонимами этих дескрипторов), причем дескрипторы должны быть систематизированы по смыслу, а смысловые связи между ними эксплицитно выражены. Именно в таком, и только в таком, значении мы будем далее употреблять термин «тезаурус».

Для построения нормативных тезаурусов исключительно большое значение имеет выявление и обозначение наиболее важных смысловых (парадигматических) связей, существующих между дескрипторами. Именно этот вопрос и будет подробно рассмотрен в следующем разделе.

## ПАРАДИГМАТИЧЕСКИЕ ОТНОШЕНИЯ МЕЖДУ ДЕСКРИПТОРАМИ

И. И. Ревзин называет парадигматическими отношениями такие, которые основаны на сходстве означающих и/или означаемых двух знаков, не находящихся в синтагматическом отношении [102]. Из этого определения следует, что к числу парадигматиче-

[101] R. R. Freeman and P. Atherton. Указ. соч., стр. 18.

[102] Как было указано ранее, синтагматические отношения основаны на линейной упорядоченности отдельных знаков в тексте.

ских отношений относятся также отношения *синонимии* (тождество означаемых при различии означающих) и отношения *омонимии* (тождество означающих при различии означаемых) [103]. Однако в данной главе, посвященной рассмотрению дескрипторных информационно-поисковых языков, термин «парадигматические отношения» употребляется в более узком значении. *Под парадигматическими отношениями здесь понимаются лишь такие отношения между словами (означающими), которые основаны на существовании тех или иных связей между означаемыми.* Таким образом, к категории парадигматических отношений мы не относим отношения омонимии, а также отношения, основанные на одинаковости основы при разных окончаниях (парадигмы склонения и спряжения) и на одинаковости окончания при разных основах. Что же касается отношений синонимии, то они рассматриваются отдельно от других парадигматических отношений. Это объясняется тем, что в само́м дескрипторном ИПЯ синонимов нет и что отношения синонимии существуют между дескрипторами и лексическими единицами соответствующего естественного языка, т. е. ключевыми словами. Таким образом, в данном случае мы имеем дело уже не с внутриязыковыми, а с межязыковыми отношениями.

Теперь попытаемся выяснить, какой смысл вкладывают в понятие «парадигматические отношения» различные специалисты в области информационного поиска. Как уже указывалось ранее, Дж. Перри, А. Кент и М. Берри называют отношения такого типа *аналитическими*. Они считают, что аналитические отношения — это отношения между понятиями и что в ИПЯ следует использовать такие виды аналитических отношений, которые обеспечивают наибольшую эффективность информационного поиска. При этом Дж. Перри, А. Кент и М. Берри допускают возможность произвольного установления того или иного аналитического отношения между двумя понятиями, если это способствует повышению эффективности информационного поиска [104].

Э. С. Бернштейн, Д. Г. Лахути и В. С. Чернявский называют примерно такой же тип отношений *базисными* и определяют их как отношения, существующие между словами поискового языка независимо от конкретных текстов [105].

У этих авторов установление базисных отношений между словами поискового языка определяется не столько фактом суще-

---

[103] И. И. Ревзин. Метод моделирования и типология славянских языков. М., «Наука», 1967, стр. 13; А. И. Уемов. Проблема синонимов и современная логика. В сб.: «Логико-грамматические очерки». М., «Высшая школа», 1961, стр. 26—48.

[104] J. W. Perry, A. Kent and M. M. Berry. Machine literature searching. New York, Interscience Publishers, 1956, стр. 77—79.

[105] См.: Э. Бернштейн, Д. Лахути и В. Чернявский. Вопросы теории поисковых систем. М., ВНИИЭМ, 1966, стр. 69.

ствования определенной связи между соответствующими означаемыми, сколько стремлением повысить эффективность информационного поиска. Излагая особенности поискового языка разработанной ими ИПС типа «Пусто-Непусто-4», Э. С. Бернштейн, Д. Г. Лахути и В. С. Чернявский пишут: «Слова поискового языка частично упорядочены развитой системой базисных отношений, которые используются в формулировке критерия смыслового соответствия. Базисные отношения системы, в основном, охватывают как родо-видовые отношения слов естественного языка, так и отношения синонимии, но не совпадают с ними. Система этих отношений строилась не умозрительно, а с учетом единственного того эффекта, который вызывался той или иной связью при эксплуатации системы. Такой подход привел к довольно существенным отклонениям от кажущейся естественной системы родо-видовых связей и отношений синонимии» [106].

Таким образом, у Э. С. Бернштейна, Д. Г. Лахути и В. С. Чернявского базисные отношения являются в значительной степени отношениями ad hoc, которые устанавливаются для конкретной ИПС и действуют только в ее пределах. Это означает, что при разработке поискового языка для любой другой отрасли науки или техники (ИПС типа «Пусто-Непусто-4» разработана для электротехники) трудоемкий процесс установления базисных отношений между дескрипторами необходимо будет каждый раз повторять заново, хотя в этот новый язык может входить фрагмент словарного состава ИПЯ по электротехнике.

Р. Крос, Ж.-К. Гардэн и Ф. Леви определяют парадигматические отношения в ИПЯ как некоторые семантические отношения, устанавливаемые a priori между словами этого языка. Одним из синонимов термина «парадигматические отношения» они считают термин «классификация» [107].

Другие французские исследователи под парадигматическими отношениями понимают «общепринятые базисные (essentiels) и вследствие этого — имплицитные отношения, которые существуют между словами» [108]. Они подразделяют парадигматические отношения на иерархические и «боковые» (latéraux). «Боковыми» отношениями между словами в данном случае называются такие симметричные отношения, которые основаны на сходстве предме-

[106] Э. С. Бернштейн, Д. Г. Лахути, В. С. Чернявский. О критериях оценки поисковых систем. Сб.: «Научно-техн. информация», 1964, № 3, стр. 29.

[107] R. C. Cros, J.-C. Gardin, F. Lévy. L'Automatisation des recherches documentaires. Un modèle général le Syntol. Paris, Gauthier-Villars, 1964, стр. 247.

[108] Les langages documentaires — Modele descriptif et problemes fondamentaux. В кн.: «Symbolic languages in data processing». (Proceedings of the Symposium organized and edited by the International Computation Center, Rome, March 26—31, 1962). New York, Science Publishers, 1962, стр. 655.

тов, обозначаемых этими словами, или на их принадлежности к одному и тому же классу.

Цель данного раздела состоит в том, чтобы выявить наиболее существенные парадигматические отношения между словами, обусловленные не какими-то сугубо практическими соображениями, а исключительно свойствами самих предметов и явлений, обозначаемых этими словами, или смысловым содержанием понятий, выражаемых ими. Такие парадигматические отношения должны быть действительны в пределах естественного языка, а также в любом дескрипторном ИПЯ, словарный состав которого представляет собой фрагмент лексики этого естественного языка.

Необходимо также разработать более или менее строгую методику построения классов дескрипторов, связанных между собой парадигматическими отношениями. Без этого процесс установления парадигматических отношений между дескрипторами в тезаурусах будет по-прежнему оставаться своеобразным искусством, основанным на сочетании содержательных суждений с интуицией и прагматическими соображениями.

Одним из основных проявлений тех или иных парадигматических отношений между словами являются ассоциации, вызываемые этими словами. Поэтому ниже кратко излагается сущность учения об ассоциациях, а также намечаются пути использования психологических методов для выявления парадигматических отношений между дескрипторами.

В реальной действительности связи между разными предметами и явлениями могут иметь такой характер, что восприятие, воспоминание или размышление об одном предмете или явлении всегда влечет за собой припоминание других предметов или явлений, которые в чем-то сходны с этим предметом (явлением) или как-то связаны с ним в пространстве и/или времени. Физиологической основой этого запоминания и воспроизведения служат временные, но достаточно прочные связи, которые образуются в коре головного мозга человека под влиянием объективных связей, существующих между разными предметами или явлениями. Такие условные связи возникают не только между предметами (их образами), но и между словами, обозначающими эти предметы. В психологии эти связи называются ассоциациями (от лат. ad — при, socio — соединяю, сочетаю) [109]. «Временная нервная связь, — отмечал И. П. Павлов, — есть универсальнейшее физиологическое явление в животном мире и в нас самих. И вместе с тем оно же и психическое — то, что психологи называют ассоциацией, будет ли это образование соединений

---

[109] Психология. Учебник для пед. институтов. Под ред. А. А. Смирнова и др. Изд. 2. М., Учпедгиз, 1962, стр. 204—207.

из всевозможных действий, впечатлений или из букв, слов или мыслей» [110].

Ассоциации образуются в процессе восприятий, в процессе воспоминаний (ассоциации представлений и суждений) и в процессе размышлений (ассоциации понятий, суждений и умозаключений).

Психологи различают два рода ассоциаций: *простые* и *сложные*. К простым ассоциациям относятся:

а) ассоциации по смежности (в пространстве и времени);
б) ассоциации по сходству;
в) ассоциации по контрасту.

Ассоциации по смежности — наиболее простой вид ассоциаций. Такие ассоциации являются отражением связей предметов и явлений в пространстве и времени. К пространственным связям относятся связи между представлениями предметов и явлений, которые воспринимались в непосредственной близости друг от друга, а к временны́м — связи между представлениями предметов и явлений, которые воспринимались в одно и то же время или непосредственно следовали одно за другим. Примеры ассоциаций по смежности: *стул — стол*, *листопад — осень*, *желудь — дуб* и т. д.

Ассоциации по сходству возникают в результате обобщения (генерализации) условной связи. Они выступают тогда, когда представления одних предметов или явлений вызывают в сознании человека представления других предметов и явлений, сходных с первыми (внешне или по какому-либо существенному признаку).

Примеры ассоциаций по сходству: *озеро — море*, *газ — туман — дым — пар*, *часы — весы — термометр* и т. д.

И. П. Павлов следующим образом охарактеризовал ассоциации по смежности и сходству: «Условная связь... есть, очевидно, то, что мы называем ассоциацией по одновременности. Генерализация условной связи отвечает тому, что зовется ассоциациями по сходству. Синтез и анализ условных рефлексов (ассоциаций) — в сущности те же основные процессы нашей умственной работы» [111].

Ассоциации по контрасту характеризуются тем, что представление одного предмета или явления может вызвать представление другого предмета или явления, обладающего по сравнению с первым противоположными свойствами. Примеры ассоциаций по контрасту: *тяжелый — легкий*, *прозрачный — непрозрачный*, *капитализм — социализм* и т. д.

---

[110] И. П. Павлов. Полное собр. соч., изд. 2, т. 3, кн. 2. М., Изд-во АН СССР, 1951, стр. 325.
[111] И. П. Павлов. Указ. соч., стр. 335.

Кроме простых ассоциаций — по смежности, сходству и контрасту, существуют также сложные ассоциации — по смыслу. Такие ассоциации возникают в тех случаях, когда между предметами и явлениями постоянная связь существует также и в самой действительности. Главными из смысловых ассоциаций, составляющих основу наших знаний, являются ассоциации *вид — род, часть — целое* и *причина — следствие* [112]. «По классификации И. П. Павлова, — пишет Е. В. Шорохова, — родовое понятие ассоциации включает несколько видов:

а) временная связь безусловного раздражителя с любым достаточной силы внешним воздействием — условным раздражителем в классическом условном рефлексе (пищевом, оборонительном и т. д.);

б) связь в действительности не связанных между собой явлений, возникающая лишь в силу их одновременного или последовательного воздействия на нервную систему животных или человека (случайные ассоциации);

в) связь явлений, одновременно действующих на нервную систему и в действительности постоянно связанных между собой. (Подчеркнуто нами. — *Авторы*).

Последний вид ассоциаций Павлов считал основой знаний, основой научного принципа каузальности, причинности, поскольку эти ассоциации отражают реально существующие связи между предметами и явлениями объективной действительности. Третий вид ассоциаций, отражающих причинно-следственные отношения, специфичен для мыслительной деятельности человека» [113].

В учении об ассоциациях нас интересует прежде всего понимание их в качестве связей, которые отражают объективно существующие отношения реальной действительности, так как именно в таких смысловых ассоциациях и находят свое проявление парадигматические отношения между словами.

Мысль о том, что в системе языка слова связаны двумя типами связей — по сходству (внутреннему или внешнему) и по смежности, когда слово рассматривается в сочетании с другими словами, с которыми оно чаще всего употребляется, высказывалась известным русским лингвистом Н. В. Крушевским еще в 80-х годах прошлого века. В своей книге «Очерк науки о языке» (Казань, 1883 г.) он писал:

«... Всякое слово связано с другими словами узами ассоциаций по сходству; это сходство будет не только внешнее, т. е. звуковое или

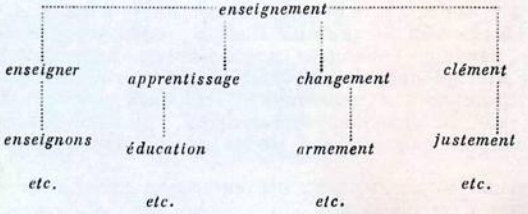
[112] Психология. Учебное пособие для пед. институтов. Под ред. А. С. Ковалева и др. М., «Просвещение», 1966, стр. 179.

[113] Е. В. Шорохова. Учение И. П. Павлова о высшей нервной деятельности и его значение для психологии мышления. В сб.: «Исследования мышления в сов. психологии». М., «Наука», 1966, стр. 88.

структурное, морфологическое, но и внутреннее, семасиологическое. Или, другими словами: всякое слово способно, вследствие особого психического закона, возбуждать в нашем духе другие слова, с которыми оно сходно, и возбуждаться этими словами. . .

<. . .> Если, вследствие закона ассоциации по сходству, слова должны укладываться в нашем уме в системы или гнезда, то благодаря закону ассоциации по смежности те же слова должны строиться в ряды»[114]. И далее: «Если представление о вещи неразлучно с представлением о соответствующем слове, то что же из этого следует? Слова должны классифицироваться в нашем уме в те же группы, что и обозначаемые ими вещи»[115].

Позже мысли о существовании ассоциативных связей между словами были развиты Ф. де Соссюром (1916 г.), который попытался графически представить эти связи следующим образом:




В приведенной схеме четыре линии ассоциаций исходят из существительного *enseignement* (преподавание): первая слева линия связывает существительное *enseignement* с глаголом *enseigner* (преподавать) на основании их формального и семантического сходства (эти слова имеют общую корневую морфему). Вторая линия связывает слово *enseignement* со словом *apprentissage* (учение) и *éducation* (образование) на основании их семантического сходства. Третья линия связывает слово *enseignement* со словами *changement* (изменение), *armement* (вооружение) и т. д. на том основании, что все эти абстрактные существительные образованы из глаголов при помощи суффикса *-ment*. Четвертая линия связывает существительное *enseignement* с прилагательным *clément*

---

[114] Цит. по: В. В. Щеулин и В. И. Медведева. Хрестоматия по истории грамматических учений в России. М., «Высшая школа», 1965, стр. 305.

[115] Там же, стр. 306.

(мягкий, милосердный) и наречием *justement* (справедливо, точно) лишь на основании случайного совпадения окончаний этих слов [116].

Ф. де Соссюр писал, что в такой схеме «любой данный термин подобен центру созвездия или точке, в которой сходится неопределенное число соподчиненных терминов (termes coordonés)» [117].

Идеи Ф. де Соссюра об «ассоциативных созвездиях» (constellation associative) получили дальнейшее развитие в работах его ученика Ш. Балли, который расценивал собственно языковые связи между словами как ассоциативные и говорил об «ассоциативных полях», объединяющих отдельные слова.

«. . . Языковая система, — писал Ш. Балли, — представляется нам в виде обширной сети постоянных мнемонических ассоциаций, весьма сходных между собой у всех говорящих субъектов, — ассоциаций, которые распространяются на все части языка от синтаксиса, стилистики, затем лексики и словообразования до звуков и основных форм произношения (ударения, интонации, продолжительности звучаний, пауз и т. п.)» [118].

В другой своей работе Ш. Балли следующим образом определил понятие «ассоциативного поля»: «Ассоциативное поле — это венец (halo), которым окружен знак и внешние края которого сливаются с окружающей их средой. . . Слово *вол* заставляет вас думать (1) о «корове, буйволе, теленке, рогах, жевании жвачки, мычании» и т. д.; (2) о «пахоте, плуге, ярме» и т. д.; наконец (3) оно может вызвать и вызывает у говорящих на французском языке представление о силе, выносливости, упорном труде, но также и о медлительности, неповоротливости и пассивности» [119].

Итак, в системе любого языка каждое слово окружено сложной сетью ассоциаций, образующей его «ассоциативное поле». Эти ассоциации связывают данное слово с другими словами, которые имеют с ним что-то общее (по смыслу или звучанию). Некоторые из этих ассоциаций основаны на смысловых связях слов, а другие имеют чисто формальное сходство. Имеются также ассоциации, основанные на смысловом и формальном сходстве (например, грамматические парадигмы глагола, существительного и т. д.).

Например, слово «свет» связано по смыслу со словами «темнота», «день», «солнце» и т. п. На основании формального сходства оно связано со словом «свет» (земля, мир), а на основании

---

[116] F. de Saussure. Cours de linguistique générale, éd. 5. Paris, 1955, стр. 175.

[117] Там же, стр. 174.

[118] Ш. Балли. Общая лингвистика и вопросы французского языка. М., ИЛ, 1955, стр. 30—31.

[119] Ch. Bally. L'arbitraire du signe. «Le Français moderne», 1940, v. 8, juillet, стр. 195.

смыслового и формального сходства — со словами «светлый», «светить», «свеча» и т. д.

Если вообразить, что все три типа ассоциаций — по смыслу, по формальному сходству, по смыслу и формальному сходству — могут как бы исходить в разных направлениях от одного слова, а также учесть, что в случае многозначности слова сеть его ассоциативных связей еще более усложнится, то можно получить некоторое представление о бесконечной сложности семантических отношений, существующих между словами естественного языка.

По определению ассоциативное поле — это открытая структура. Некоторые из образующих такое поле ассоциаций являются субъективными, тогда как их основная часть одинакова для большинства говорящих. Для идентификации основных видов ассоциаций между словами был проведен ряд психологических экспериментов [120]. (Несколько подробней об этом будет сказано ниже.) Однако такую задачу можно решить и путем использования чисто лингвистических методов (например, путем подбора наиболее явных синонимов и омонимов какого-либо слова; путем подбора других слов, сходных с данным словом по звучанию и смыслу, и т. д.).

Нам также нужно из множества разных ассоциативных отношений выбрать наиболее существенные, т. е. такие, которые имеют в процессах мышления наибольший «удельный вес». Затем путем анализа отобранных отношений мы попытаемся составить перечень отношений, принадлежащих к категории парадигматических (в указанном ранее смысле). Для решения первой части этой задачи — отбора наиболее существенных ассоциативных отношений — воспользуемся результатами соответствующих психологических экспериментов.

Известно, что одни и те же раздражители могут вызывать разные виды ассоциаций. Некоторые из этих ассоциаций закрепляются и становятся значительно более прочными, чем другие. При появлении тех же раздражителей более прочные ассоциации доминируют и проявляются раньше. Основным способом количественной оценки сравнительной прочности (силы) разных видов ассоциаций является определение частоты совместной встречаемости разных слов.

В 1910 г. Дж. Кент и А. Розанов провели такой ассоциативный эксперимент. Был составлен список из 100 слов-раздражителей (существительных и прилагательных). Испытуемый, сидящий спиной к экспериментатору, должен был называть первое слово, приходящее ему на ум в ответ на названное эксперимен-

---

[120] T. Cazacu. La «structuration» dynamique des significations. В кн.: «Mélanges linguistiques publiés à l'occasion du VIII-e Congrès International des Linguistes». Bucharest, 1957, стр. 113—127.

татором слово-раздражитель. Такому испытанию были подвергнуты 1000 взрослых мужчин и женщин, имеющих разное образование и разные профессии, причем каждому испытуемому назывались все 100 слов-раздражителей. В 1916 г. Г. Вудро и Ф. Ловелл систематизировали результаты этого эксперимента (вместе с результатами своего эксперимента с 1000 детей).

Классифицировались лишь реакции, возникавшие не менее чем у 10 испытуемых, т. е. наиболее частые реакции. Они составили около 80% всех реакций. Эти реакции были распределены по 21 классу. Результаты данной классификации приведены в табл. 24. Очевидно, что для отнесения ответных реакций к классам, перечисленным в этой таблице, пришлось принимать много произвольных решений. Кроме того, ответные реакции в значительной степени зависели от выбора слов-раздражителей и испытуемых. Поэтому приведенные в таблице цифры можно рассматривать лишь как приблизительные оценки силы ассоциаций разных типов [121].

В связи с этим необходимо отметить, что классификация реакций при свободных ассоциациях всегда условна, поскольку иногда бывает весьма трудно отнести к тому или иному классу даже часто встречающиеся ассоциации. Это можно видеть на следующих примерах [122]:

| | |
|---|---|
| Человек — животное: | суперординация или координация? |
| Рука — нога: | координация, контраст или сходство? |
| Красный — цвет: | суперординация или прилагательное-существительное? |
| Соленый — сладкий: | координация или контраст? |

Если исходить лишь из свободных ассоциаций, то ответ на любой из этих вопросов будет зависеть от точки зрения испытуемого, а не от отношений между понятиями, выражаемыми этими словами. Однако нас интересуют именно отношения между понятиями (означаемыми).

Рассмотрим теперь табл. 24 и сделаем следующее:

а) исключим из нее все данные об ассоциациях у детей;

б) объединим в общие классы ассоциации вида (4) и (12); (5), (8) и (18); (10) и (13); (9) и (17);

в) все остальные виды ассоциаций, относительная частота которых не превышает 2,0%, включим в класс «Разные».

В результате такого преобразования мы получим табл. 25. В этой таблице осталось всего 9 видов ассоциаций. Из этих

---

[121] G. A. Miller. Language and communication. New York, McGraw—Hill, 1963, стр. 179—180.

[122] Р. Вудвортс. Экспериментальная психология. М., ИЛ, 1950, стр. 575.

Таблица 24

**Относительная частота разных видов ассоциаций слов у взрослых и детей**

| Порядковый номер | Вид ассоциации | Примеры | Относительная частота, % | |
|---|---|---|---|---|
| | | | взрослые | дети |
| 1 | Координация (соподчинение) . . . . . | Стол — стул; глубокий — низкий; дом — сарай | 10,9 | 6,0 |
| 2 | Контраст . . . . . | Темный — светлый; болезнь — здоровье; глубокий — мелкий | 10,6 | 1,3 |
| 3 | Сходство . . . . . | Темный — черный; болезнь — недомогание; гора — холм | 8,9 | 8,6 |
| 4 | Суперординация (вид — род) . . . . . | Стол — мебель; музыка — звук; дом — здание | 7,6 | 3,7 |
| 5 | Прилагательное — существительное . . . . . | Глубокий — яма; мягкий — постель; черный — одежда | 6,9 | 11,2 |
| 6 | Глаголы (предикация) . . | Стол — есть; темнота — видеть; музыка — играть | 6,4 | 10,2 |
| 7 | Смежность . . . . . | Стол — посуда; темнота — ночь; болезнь — врач | 6,0 | 15,3 |
| 8 | Существительное — прилагательное . . . . . | Музыка — приятная; гора — высокая; дом — большой | 4,3 | 7,8 |
| 9 | Причина — следствие . . | Болезнь — смерть; лампа — свет; купание — чистый | 2,5 | 1,9 |
| 10 | Целое — часть . . . . . | Стол — ножка; музыка — ноты; гора — скалы | 2,1 | 3,6 |
| 11 | Причастия . . . . . | Музыка — поющий; едящий — пьющий; стул — сидящий | 1,9 | 0,9 |
| 12 | Субординация (род — вид) | Музыка — песня; болезнь — лихорадка; фрукты — яблоко | 1,6 | 2,1 |
| 13 | Часть — целое | Плод — дерево; ступня — нога; солдат — армия | 1,1 | 0,4 |

Таблица 24 (*продолжение*)

| Порядковый номер | Вид ассоциации | Примеры | Относительная частота, % | |
|---|---|---|---|---|
| | | | взрослые | дети |
| 14 | Материал . . . . . | Стол — дерево; игла — сталь; лампа — стекло | 1,0 | 0,9 |
| 15 | Глагол (действие) — объект | Есть — хлеб; (пилить — доску); вбивать — гвозди | 0,9 | 1,7 |
| 16 | Завершение . . . . . | (Вперед — марш); стол — накрыть; паук — паутина | 0,77 | 1,04 |
| 17 | Следствие — причина . . . | Сон — уставший; аппетит — голодный; бояться — грабитель | 0,4 | 0,5 |
| 18 | Существительное — абстрактный атрибут . . . . | Гора — высота; бабочка — красота; орел — полет | 0,25 | 0,05 |
| 19 | Ассонанс . . . . . | (Борт — порт; карта — парта; мел — сел) | 0,07 | 0,43 |
| 20 | Местоимения . . . . . | Женщина — она; желудок — мой; жажда — у меня | 0,05 | 0,22 |
| 21 | Разные . . . . . | | 4,7 | 5,5 |
| | Сумма . . . | | 78,94 | 83,34 |

ассоциаций нам нужно выбрать такие, которые обусловлены существованием между словами отношений парадигматического типа. В связи с этим полезно напомнить, что парадигматическими мы условились считать отношения, основанные на существовании объективных связей между означаемыми, т. е. между внутренними, семантическими сторонами слов. Таким образом, парадигматическим отношениям соответствуют лишь такие виды ассоциаций, члены которых связаны между собой семантическими (смысловыми) отношениями.

Термин «семантическая связь» нуждается в некотором уточнении. Семантической или смысловой мы будем называть такую, и только такую связь, которая существует между словами или понятиями, а не между предметами, фактами. Это уточнение позволит нам как-то ориентироваться в бесконечно сложной системе всеобщих связей и взаимодействий, которые существуют в объективном мире.

Например, между словами «птица» и «летать» существует определенная связь, которая обусловлена тем фактом, что птицы

Эти виды ассоциаций составляют не менее $^1/_3$ всех свободных ассоциаций, которые были идентифицированы при проведении описанных выше экспериментов.

Полезно отметить, что составленный нами список смысловых ассоциаций в значительной степени совпадает с перечнем так называемых внутренних ассоциаций М. Траутшольда — автора первого экспериментального исследования ассоциаций (1883 г.). М. Траутшольд делил все ассоциации на внешние и внутренние. Внешними он считал ассоциации, возникающие только потому, что предметы представлений один или несколько раз воспринимаются нами либо в один и тот же момент времени, либо в моменты, непосредственно следующие один за другим. Напротив, внутренние ассоциации возникают в силу того, что предметы представлений имеют что-то общее, чем-то родственны друг другу. Члены внутренних ассоциаций связаны друг с другом отношениями вид — род, соподчинения, причина — действие, цель — средство [127].

Ч. Берньер и К. Хьюманн считали, что для информационного поиска наиболее [illegible] ными аналитическими отношениями между «семант[illegible] ментами смысла, дескрипторами) являются след[illegible] ые отношения:

— часть — целое,
— математико-функциональное,
— определительное.

Под опр[illegible] м отношением Ч. Берньер и К. Хьюманн понимают т[illegible] ение, которое уточняет смысловое содержание деск[illegible] Например, определительное отношение существует ме[illegible] скрипторами «стекло» и «стекловидные материалы», «цемен[illegible] «бетон», «продовольствие» и «питание» и т. д.[128]

Дж. Перри и А. Кент использовали в своем Семантическом коде 10 следующих видов аналитических отношений:

— категорическое (вид — род),
— включения (целое — часть),
— вхождение (часть — целое),
— объединения,
— продуктивное,
— аффективное,
— инструментальное,
— отрицания,
— атрибутивное,
— сходства.

Эти аналитические отношения здесь перечислены лишь для того, чтобы их можно было сопоставить с парадигматическими

[127] См.: П. А. Шеварев. Обобщенные ассоциации в учебной работе школьника. М., Изд-во Академии пед. наук РСФСР, 1959, стр. 31—32.

[128] C. L. Bernier and K. F. Heumann. Correlative indexes. III. Semantic relations among semantemes. The technical thesaurus «American Documentation», 1957, v. 8, № 3, 212; C. L. Bernier. Correlative indexes. IX. Vocabulary control. «Journal of Chemical Documentation», 1964, v. 4, № 2, стр. 99.

отношениями, которые были отобраны нами путем анализа разных видов ассоциаций. Такое сопоставление показывает, что, по крайней мере, шесть из этих десяти аналитических отношений полностью или частично совпадают с нашими парадигматическими отношениями (имеются в виду отношения: *категорическое*, *включения*, *вхождения*, *продуктивное*, *инструментальное* и *сходства*)[129].

Примерно такую же картину мы получаем и при сопоставлении наших парадигматических отношений с аналитическими отношениями, которые были постулированы Ч. Берньером и К. Хьюманном. Результаты этих сопоставлений могут служить дополнительным свидетельством того, что отбор наиболее важных парадигматических отношений был произведен в целом правильно.

Составленный нами список парадигматических отношений был использован для разработки алгоритма, который по любому дескриптору порождает для тезауруса словарную статью этого дескриптора. Под словарной статьей тезауруса в данном случае понимается упорядоченный класс дескрипторов, связанных парадигматическими отношениями с «заглавным» дескриптором, который обозначает этот класс. В словарную статью включаются также все ключевые слова, которые в пределах данной ИПС считаются синонимами «заглавного» дескриптора. Словарная статья тезауруса может быть выполнена в виде списка, таблицы и/или графической схемы, в которых при помощи соответствующих помет, шифров, стрелок и т. п. обозначены виды парадигматических отношений, связывающих «заглавный» дескриптор с другими перечисленными в его словарной статье дескрипторами.

Как будет показано ниже, в словарных статьях тезаурусов все дескрипторы, связанные с «заглавным» дескриптором теми или иными парадигматическими отношениями, обычно даются в виде общего алфавитного списка. Исключение составляют лишь дескрипторы, которые связаны с «заглавным» дескриптором отношениями «род — вид» и «вид — род». Такие дескрипторы группируются в словарных статьях в отдельные подклассы, также упорядоченные по алфавиту. В связи с этим может возникнуть вопрос: нельзя ли вообще ограничиться выделением лишь двух типов парадигматических отношений — «род — вид» («вид — род») и «прочие парадигматические отношения». Ведь в словарных статьях тезаурусов эксплицитно выражаются только эти два типа парадигматических отношений (если не считать отношений синонимии между дескрипторами и ключевыми словами).

На такой вопрос мы даем отрицательный ответ, так как *без списка важнейших парадигматических отношений невозможно*

---

[129] См.: B. C. Vickery. On retrieval system theory. 2nd ed. London, Butterworths Scientific Publications, 1965, стр. 40.

*сформулировать более или менее однозначные правила построения подкласса «Другие ассоциативные дескрипторы» в словарных статьях тезаурусов.*

Правда, имеются и такие тезаурусы, в словарные статьи которых включаются все дескрипторы, связанные с «заглавным» дескриптором любыми ассоциативными связями. Эти словарные статьи порождаются путем статистической обработки поисковых образов документов и поисковых предписаний, причем в словарной статье против каждого дескриптора указывается (например, в процентах) частота его совместной встречаемости (ассоциации) с «заглавным» дескриптором. Тезаурусы такого типа относятся к классу статистических тезаурусов и называются *ассоциативными матрицами слов.* Идея использования ассоциативных матриц ключевых слов в качестве упрощенных тезаурусов принадлежит Дж. Костелло и Ю. Уоллу (1959 г.) [130]. Ниже в качестве примера приводится фрагмент ассоциативной матрицы слов [131].

| | |
|---|---|
| АБАК | |
| СУММИРУЮЩАЯ МАШИНА | 0,65 |
| СЧЕТНО-РЕШАЮЩАЯ МАШИНА | 0,35 |
| ВЫЧИСЛИТЕЛЬНАЯ МАШИНА | 0,10 |
| КИТАЙСКИЙ | 0,05 |
| ВЫЧИСЛИТЕЛЬНАЯ МАШИНА | |
| АНАЛОГОВЫЙ | 0,30 |
| ЦИФРОВОЙ | 0,25 |
| АБАК | 0,10 |
| ЭЛЕКТРОННЫЙ | |
| ЦИФРОВОЙ | 0,70 |
| АНАЛОГОВЫЙ | 0,20 |
| АППАРАТ | 0,05 |

Основное достоинство статистических тезаурусов типа ассоциативных матриц слов заключается в том, что их можно строить при помощи счетно-перфорационных и вычислительных машин.

Рассмотрим теперь отдельно парадигматические отношения «род — вид» («вид — род»), так как этот вид отношений играет особенно важную роль при информационном поиске. Кроме того,

---

[130] J. C. Costello, Jr. and E. Wall. Recent improvements in techniques for storing and retrieving information. В кн.: M. Taube (comp.): Emerging solutions for mechanizing the storage and retrieval of information. Washington, Documentation, 1959, стр. 141—169. (Studies in coordinate indexing, v. 5).

[131] W. F. Williams. Principles of automated information retrieval. Elmhurst, Ill., The Business Press, 1965, стр. 262.

*через родо-видовые отношения реализуется еще один важный вид парадигматических отношений — отношение соподчинения.* Например, между дескрипторами ФТОР, ХЛОР, БРОМ и ИОД устанавливается соподчинительная (координативная) связь, если каждый из этих дескрипторов будет связан отношением «вид — род» с дескриптором ГАЛОГЕНЫ.

В словарном составе ИПЯ встречается немало дескрипторов, выражающих такие понятия, которые полностью включают в себя одно или несколько более узких понятий, выражаемых другими дескрипторами. Например, понятие «сплавы» полностью включает в себя понятие «сплавы меди», а это понятие, в свою очередь, полностью включает в себя понятия «латунь» и «бронза». Между дескрипторами, выражающими такие понятия, объективно существуют отношения рода и вида. Выявление родовидовых отношений между дескрипторами обычно производится на основе сопоставления объемов понятий, выражаемых этими дескрипторами.

*Понятие* есть мысль, в которой отражаются и фиксируются наиболее существенные свойства, признаки и связи предметов и явлений объективной действительности. Ф. Энгельс подчеркивает, что «результаты, в которых обобщаются данные его (естествознания — *авторы*) опыта, суть понятия . . .» [132]. Совокупность свойств, признаков и отношений, отражаемых в понятии, называется *содержанием понятия*, а класс (множество) предметов, относящихся к содержанию понятия, — *объемом понятия* [133]. Если уменьшить содержание понятия на один признак, то получится новое понятие, имеющее больший объем и связанное с исходным понятием отношением «род—вид».

Итак, термины «вид» и «род» применяются для обозначения двух классов, из которых первый составляет правильную (не совпадающую с целым) часть второго класса и объединяет предметы со специфическими для них свойствами. В связи с этим определением необходимо указать на недопустимость отождествления отношений «вид—род» и «часть—целое», хотя между ними и существует определенная связь. «Вид—род» есть отношения между классами, тогда как «часть—целое» — это отношение между предметами. Земля — есть часть Солнечной системы, но не вид ее. Е. К. Войшвилло пишет по этому поводу следующее: «Род есть только совокупность видов и притом совокупность, не представляющая собой реального предмета (как и каждый класс, являющийся видом данного рода, есть лишь совокупность подклассов — видов данного класса, а в конечном счете — индивидов). Следовательно, род в противопоставлении видам (как и вид

---

[132] К. Маркс и Ф. Энгельс. Сочинения, т. 20, стр. 14.
[133] Д. П. Горский. Логика. М., Учпедгиз, 1965, стр. 41—42.

в противопоставлении индивидам) есть лишь абстракция. Между тем предмет, хотя и состоит из частей, сам представляет собой реальную (существующую в пространстве и времени) вещь; таковыми же являются и его части» [134].

Понятие, отражающее существенные признаки класса предметов, в состав которого входят другие классы предметов, являющиеся видами этого рода, называется *родовым понятием*. Например, понятие «перфокарта» есть родовое по отношению к понятию «суперпозиционная перфокарта». Соответственно *видовым* называется такое понятие, которое отражает существенные признаки класса предметов, являющегося видом какого-либо класса предметов и входящего в состав этого класса.

Одно и то же понятие одновременно может быть как видовым, так и родовым в зависимости от того, по отношению к какому другому понятию оно рассматривается. Например, понятие «прямоугольник» является видовым по отношению к понятию «параллелограмм» и родовым по отношению к понятию «квадрат». Исключение составляют лишь категории и единичные понятия. *Категория* — это предельно широкое понятие, в котором отражены наиболее общие свойства, связи и отношения предметов и явлений объективного мира (например, «материя», «пространство», «время» и т. д.). Для категорий не может существовать родовых понятий, так как категории уже не входят в более широкие понятия. *Единичное понятие* — это такое понятие, в котором отражены существенные и отличительные признаки какого-либо одного предмета или явления (например, «классификация Дьюи», «селекторный код» и т. д.). Для единичных понятий не существует понятий видовых. Взаимосвязь видовых и родовых понятий отображает объективно существующие отношения вида и рода в природе и обществе.

По отношению к видовому понятию родовое понятие является *подчиняющим*, так как его объем включает в себя — в качестве своей части — объем видового понятия. Соответственно видовое понятие по отношению к родовому является *подчиненным*. Понятия, которые в равной мере подчинены одному родовому понятию, называются *соподчиненными*. Например, понятия «фонетика», «лексикология» и «грамматика» являются понятиями, соподчиненными одному родовому понятию «языкознание». Родовидовые отношения между понятиями, как уже говорилось выше, удобно изображать при помощи кругов Эйлера, а также в виде графов, называемых деревьями.

Для каждого дескриптора в словарном составе ИПЯ должны быть найдены соответствующие родовые и видовые дескрипторы. При этом следует по возможности воздерживаться от введения

---

[134] Е. К. Войшвилло. Понятие. М., Изд-во МГУ, 1967, стр. 170.

в словарный состав ИПЯ отсутствующих в нем родовых и видовых дескрипторов. В противном случае это может привести к значительному расширению словарного состава ИПЯ, к его загромождению редко используемыми дескрипторами, что существенно затруднит процедуру перевода с естественного языка на этот ИПЯ.

При отыскании родовых и видовых дескрипторов в словарном составе ИПЯ совершенно недопустима какая-либо подгонка содержания понятий, выражаемых соответствующими дескрипторами, под структуру предварительно составленной классификационной системы, т. е. установление каких-либо искусственных родо-видовых отношений между дескрипторами. Многие дескрипторы могут вообще не иметь в ИПЯ родовых и/или видовых дескрипторов. Но это не должно рассматриваться как недостаток этого ИПЯ.

В словарном составе ИПЯ встречаются дескрипторы, которые выражают очень широкие понятия. Такие дескрипторы целесообразно снабдить пояснительными пометами, которые бы в нужной степени блокировали родо-видовые связи этих дескрипторов с другими дескрипторами данного ИПЯ. Например, дескриптор «химические элементы» обязательно должен быть снабжен ограничительной пометой, запрещающей автоматическое дополнение этим дескриптором поисковых образов документов, в которых говорится об отдельных химических элементах, или поисковых предписаний, которые формулируются для поиска документов о каком-либо конкретном химическом элементе. Иногда может оказаться удобным вообще исключить такой дескриптор из словарного состава ИПЯ.

Итак, для каждого дескриптора ИПЯ необходимо построить классификационное дерево, в котором были бы отображены родо-видовые связи этого дескриптора с другими дескрипторами данного ИПЯ. Предварительно должны быть сформулированы какие-то общие критерии, определяющие число уровней иерархий в родо-видовых деревьях различных семейств понятий (дескрипторов).

Весьма ценным дополнительным результатом работы по установлению родо-видовых отношений между дескрипторами является окончательное устранение неоднозначности отдельных дескрипторов, а также существенное улучшение словарного состава ИПЯ благодаря исключению из него дескрипторов, выражающих слишком широкие или слишком узкие понятия. В связи с этим необходимо отметить, что работа по установлению родо-видовых отношений между дескрипторами, как и вообще вся работа по составлению любого тезауруса, не может быть выполнена без самого активного участия в ней высококвалифицированных специалистов той отрасли знания, для которой составляется этот тезаурус.

Как показывает зарубежный опыт, при составлении тезаурусов в словарные статьи дескрипторов могут включаться даже такие родовые и видовые дескрипторы, которые в классификационном дереве находятся на расстоянии трех шагов от «заглавного» дескриптора. Вопрос о том, какие из дескрипторов, стоящих выше или ниже данного дескриптора, следует включить в словарную статью этого дескриптора, обычно решается для каждого случая отдельно. При этом учитывается содержание и объем понятия, выражаемого каждым таким дескриптором; предметная область, для которой составляется тезаурус; практические требования, предъявляемые к ИПС; эксплуатационные особенности этой ИПС и другие факторы. Однако чаще всего в словарные статьи тезауруса включаются лишь дескрипторы, стоящие непосредственно выше или ниже «заглавного» дескриптора.

В качестве примера на рис. 47 приведены классификационные деревья, построенные для ряда «заглавных» дескрипторов в «Тезаурусе ASTIA» [135]. В каждом классификационном дереве рамкой выделен «заглавный» дескриптор.

В само́м «Тезаурусе ASTIA» родо-видовые отношения каждого «заглавного» дескриптора с другими дескрипторами выражаются не графически, а в виде соответствующих записей с пометами. При этом не делается никакого различия между дескрипторами, стоящими на один, два или три шага выше (или ниже) «заглавного» дескриптора. Все дескрипторы, стоящие в классификационном дереве выше или ниже «заглавного» дескриптора, считаются соответственно его родовыми или видовыми дескрипторами. Это можно видеть на следующих примерах, которые полезно сопоставить с приведенными на рис. 47 классификационными деревьями этих же «заглавных» дескрипторов [136].

а) ХИМИЯ

Видовые дескрипторы:

МИКРОХИМИЯ

РАДИОХИМИЯ
СТЕРЕОХИМИЯ
ТЕРМОХИМИЯ
ФИЗИЧЕСКАЯ ХИМИЯ
ХИМИЧЕСКИЙ АНАЛИЗ
ЭЛЕКТРОХИМИЯ

в) ПОЛЯРОГРАФИЧЕСКИЙ АНАЛИЗ

Родовые дескрипторы:

КОЛИЧЕСТВЕННЫЙ АНАЛИЗ
ХИМИЧЕСКИЙ АНАЛИЗ

---

[135] Thesaurus of ASTIA Descriptors, 2nd ed., Arlington, Armed Services Technical Information Agency, 1962, стр. 149, 151—153, 425.

[136] Порядок расположения дескрипторов под заголовками «Родовые дескрипторы» и «Видовые дескрипторы» — алфавитный.

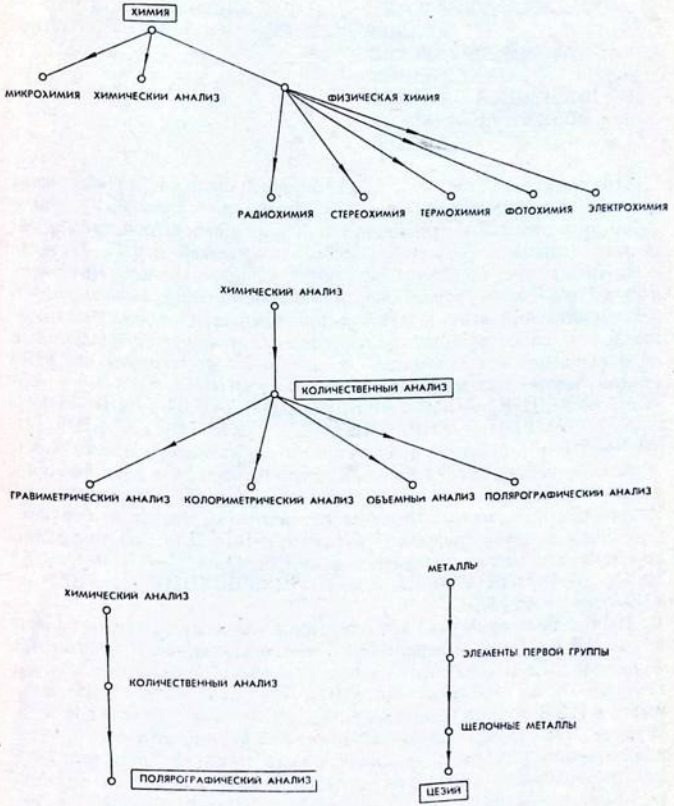



Рис. 47. Родо-видовые деревья «заглавных» дескрипторов ХИМИЯ, КОЛИЧЕСТВЕННЫЙ АНАЛИЗ, ПОЛЯРОГРАФИЧЕСКИЙ АНАЛИЗ и ЦЕЗИЙ в «Тезаурусе ASTIA»

## СПОСОБЫ ВЫРАЖЕНИЯ ПАРАДИГМАТИЧЕСКИХ ОТНОШЕНИЙ В ТЕЗАУРУСАХ

В тезаурусе парадигматические отношения между дескрипторами могут быть выражены четырьмя способами: лексикографически (т. е. при помощи системы указательных помет), при помощи таблиц, аналитически и графически.

В табл. 27 даны основные виды указательных помет, применяемых составителями тезаурусов для выражения парадигматических отношений между дескрипторами и ключевыми словами. Там же для сравнения приведены пометы, которые обычно используются в библиотечно-библиографических классификациях.

Вторым способом эксплицитного выражения в тезаурусе парадигматических отношений между дескрипторами является *табличный способ*. При использовании этого способа для выражения отношения «род-вид» между двумя дескрипторами или отношения включения между двумя классами видовой дескриптор (или имя класса) располагается под родовым дескриптором (или именем включающего класса) со сдвигом на фиксированное число знаков (обычно — на две буквы) вправо относительно первой буквы родового дескриптора (или имени включающего класса). Табличный способ выражения отношений включения (вхождения) между классами широко используется в библиотечно-библиографических классификациях.

В качестве примера такого способа выражения родо-видовых отношений между дескрипторами можно использовать предметный класс 29 — ИНФОРМАЦИЯ, ДОКУМЕНТАЦИЯ и АВТОМАТИЗАЦИЯ из «Тезауруса ASTIA», который был приведен в табл. 26. Ниже дается пример табличного расположения имен классов в фасетной классификации [137].

| | |
|---|---|
| Компоненты (реактора) | Отражатели |
| Источник нейтронов | Защитные экраны |
| Топливный элемент | Системы управления |
| Блоки | Управляющие стержни |
| Стержни | Система охлаждения, система |
| Замедлитель | теплопередачи |
| Решетка | Теплоноситель |

Такой способ позволяет выражать только родо-видовые отношения между дескрипторами или отношения включения (вхождения) между классами. Поэтому он применяется, как правило, в сочетании с другими способами и прежде всего — с лексикографическим.

Однако возможен и другой вариант табличного способа выражения в тезаурусе парадигматических отношений между дескрипторами и ключевыми словами. При использовании этого способа

[137] B. C. Vickery. Classification and indexing in science. 2nd ed. London, Butterworths Scientific Publications, 1959, стр. 204.

Таблица 27

**Основные виды указательных помет, применяемых в разных тезаурусах для обозначения парадигматических отношений между дескрипторами и ключевыми словами**

| Тезаурусы | Отношение синонимии | | Родо-видовые связи (относительно «заглавного» дескриптора) | | Прочие парадигматические отношения |
|---|---|---|---|---|---|
| | быть синонимом | иметь синонимами | родовые дескрипторы | видовые дескрипторы | |
| Тезаурус ASTIA, изд. 2 (1962 г., США) | use | includes | specific to | generic to | also see |
| Химико-технологический тезаурус (1961 г., США) | see | seen from (SF) | post on (PO) | generic to (GT) | related term (RT) |
| Тезаурус Бюро мелиорации (1963 г., США) | use | — | — | — | see also |
| Тезаурус AED (1963 г., ФРГ) | see | — | — | — | see also |
| Тезаурус технических терминов (1964 г., США) | use | used for (UF) | broader term (BT) | narrower term (NT) | related term (RT) |
| Тезаурус Бюро кораблестроения и ремонта, изд. 2 (1965 г. США) | use | includes | broader term (BT) | narrower term (NT) | related term (RT) |
| «Тезаурус Евратома», изд. 2 (1966 г.) | use, use . . . + see, see . . . or | | Выражены графически | | |
| Нормативный список предметных заголовков Американского нефтяного института | use | used for | broader terms (autoposted) | narrower terms | see also |
| Библиотечно-библиографические классификации | see | see | see also | see also | see also |

Обозначения:

| | | | |
|---|---|---|---|
| also see | — также смотри | related term (RT) | — родственный термин |
| broader term (BT) | — более широкий (по значению) термин | see | — смотри . . . |
| broader terms (autoposted) | — более широкие (по значению) термины (приписываются к «заглавному» дескриптору автоматически) | see also | — смотри также . . . |
| | | see from (SF) | — смотри от . . . |
| generic to (GT) | — родовой по отношению к . . . (о «заглавном» дескрипторе) | see or | — смотри . . . или |
| | | specific to | — видовой по отношению к . . . (о «заглавном» дескрипторе) |
| includes | — включает . . . | use | — используй, употребляй . . . |
| narrower term (NT) | — более узкий (по значению) термин | use for (UF) | — используется вместо . . . (о «заглавном» дескрипторе) |
| post on (PO) | — приписать к («заглавному» дескриптору) | | |

| I | II | III | IV | V | VI |
|---|---|---|---|---|---|
| ЖУРНАЛ-4 | | Бюллетень | Публикация | Статья<br>Журнал по документации<br>Журнал реферативный | Содержание |
| ЖУРНАЛ ПО ДОКУМЕНТАЦИИ-4 | | | Журнал | | Исследования в области документации<br>Документация документации<br>Методология |
| ЖУРНАЛ РЕФЕРАТИВНЫЙ-4 | | | Журнал | | Аннотация<br>Реферат библиографический<br>Документ вторичный<br>Карточка документационная |
| Резюме аналитическое | РЕФЕРАТ БИБЛИОГРАФИЧЕСКИЙ-8 | | | | |
| РЕЗЮМЕ АВТОРСКОЕ-3 | | | Синопсис | Вторичный документ | Статья<br>Реферат библиографический |

Рис. 48. Страница из «Тезауруса для координатного индексирования [материалов] по документации и библиотековедению» (СРР, 1965 г.)

тезаурус (или, по крайней мере, определенный его раздел) выполняется в виде таблицы, в первой вертикальной графе которой выписываются в общем алфавитном порядке дескрипторы и ключевые слова, а во второй и третьей графах — соответственно дескрипторы или ключевые слова, связанные отношением синонимии с лексическими единицами, которые стоят против них в первой графе. В последующих графах таблицы приводятся дескрипторы, связанные со стоящим в первой графе «заглавным» дескриптором родо-видовыми и другими парадигматическими отношениями.

Такой вариант табличного способа выражения парадигматических отношений (наряду с графическим способом) применен в «Тезаурусе для координатного индексирования [материалов] по документации и библиотековедению», который был опубликован Институтом технической документации СРР в 1965 г. На рис. 48 приведена страница из этого тезауруса [138].

В графе I этого тезауруса выписаны в общем алфавитном порядке все дескрипторы и ключевые слова. Дескрипторы в графах I и II даны прописными буквами и снабжены номерами предметных классов, к которым они принадлежат. Сами предметные классы выполнены в виде графических схем (см. рис. 51). В графе II приведены дескрипторы, синонимами которых считаются ключевые слова, выписанные в той же строке графы I. В графе III даны ключевые слова, считающиеся синонимами «заглавных» дескрипторов. В графах IV и V выписаны соответственно видовые и родовые дескрипторы, а в графе VI — дескрипторы, связанные с «заглавным» дескриптором другими парадигматическими отношениями.

Третий способ эксплицитного выражения в тезаурусе парадигматических отношений между дескрипторами — *аналитический*. Под аналитическим способом в данном случае понимается выражение парадигматических отношений при помощи структуры кодов (шифров) дескрипторов. Аналитический способ выражения родо-видовых отношений широко используется в библиотечно-библиографических классификациях. Например, в УДК индекс 621.375 (Усилители) означает, что класс «Усилители» является подразделением класса «Техника электрических и электромагнитных колебаний» (621.37), входящего в более широкий класс «Электротехника» (621.3); этот класс, в свою очередь, является подразделением класса «Общее машиностроение. Машиноведение» (621), входящего в еще более широкий класс «Инженерное дело. Техника» (62), который сам является подклассом класса «Прикладные науки. Медицина. Техника» (6).

Для аналитического выражения парадигматических отношений между дескрипторами применяются два типа цифровых, буквенно-цифровых или буквенных кодов — *родовые коды* (generic codes) и *семантические коды*. Родовые коды, как следует уже из их названия, предназначены для выражения лишь родо-видовых отношений между дескрипторами или отношений вхождения (включения) между классами, а семантические коды — для выражения более широкого ассортимента парадигматических отношений между дескрипторами.

Родовые коды можно называть и *связанными кодами*, так как в кодах этого типа значение каждого последующего символа (буквы или цифры) полностью определяется всеми предшествующими ему символами (если таковые имеются).

[138] Tezaur pentru indexarea coordonată în domeniul documentării şi biblioteconomiei. Bucureşti, Institutul de documentare tehnică, 1965, p. 81.

Хорошим примером связанного кода может быть индекс УДК. Как уже отмечалось, в индексе УДК любая цифра (кроме первой) не имеет собственного значения: оно определяется всеми другими цифрами, стоящими в индексе перед этой цифрой. Свободные цифры ряда 0—9 в УДК обозначают лишь главные классы предметного фасета. Цифра 5 сообщает индексу УДК 621.375 значение «Усилители» лишь тогда, когда перед ней стоят цифры 621.37. Если бы цифре 5 предшествовали, например, цифры 621.35 или 621.39, то эта цифра сообщала бы индексам 621.355 и 621.395 соответственно значения «Вторичные элементы. Аккумуляторы» и «Техника телефонной связи. Телефония». В свободном виде цифра 5 в УДК означает класс «Математика. Естественные науки».

Связанные цифровые коды применены для выражения родовидовых отношений между дескрипторами в «Тезаурусе по документации» Г. Бунтрока [139]. Ниже в качестве примера приведен фрагмент из этого тезауруса.

| | |
|---|---|
| Автоматический ввод | 3.1.2.2.1.4 |
| Автоматический вывод | 3.1.2.2.4 |
| Автоматический поиск | 3.1.2.2.3.4 |
| Автоматы | 1.2 |
| Авторы | 3.3.1.2.2 |

В этом примере шифр 3.1.2.2.3.4 *Автоматический поиск* выражает следующую иерархию классов:

| | |
|---|---|
| 3 | Обработка документов и информации |
| 3.1 | Документационная деятельность |
| 3.1.2 | Сбор, хранение, поиск |
| 3.1.2.2 | Обработка извлеченной информации |
| 3.1.2.2.3 | Поиск |
| 3.1.2.2.3.4 | Автоматический поиск |

Другой тип родовых кодов образуют такие, в которых код каждого видового дескриптора или класса образуется путем приписывания справа к коду родового дескриптора (или включающего класса) определенной *группы символов* (букв или цифр). Очевидно, что приписываемые группы символов также не имеют собственных значений. Ниже приведен пример буквенного кода такого типа [140].

| | |
|---|---|
| Животное | ЖИ |
| Млекопитающее | ЖИ МЛ |
| Собака | ЖИ МЛ СО |
| Терьер | ЖИ МЛ СО ТЕ |
| Гончая | ЖИ МЛ СО ГО |
| Охотничья собака | ЖИ МЛ СО ОХ |
| Колли | ЖИ МЛ СО КО |

[139] H. Buntrock. Thesaurus of documentation. (Provisional descriptor list). Ispra, EURATOM/CETIS, 1964.

[140] J. W. Perry, A. Kent and M. M. Berry. Machine literature searching. New York, Interscience Publishers, 1956, стр. 57

Цифровой код этого типа был применен в одной из классификаций, которая была разработана в Патентном управлении США. Этот код имеет следующий вид [141]:

| | |
|---|---|
| Гетероциклические соединения | 1313 |
| Сульфоокись паранормального бензола | 1313-1512 |
| Азолы | 1313-2512 |
| Тиазолы | 1313-2512-1423 |
| Ядро антрона | 1313-2512-1423-17 |
| Оксазолы | 1313-2512-1523 |

Семантический код основан на использовании кодов некоторого сравнительно небольшого числа элементарных понятий — *семантических множителей*, из которых по определенным правилам строятся коды более сложных понятий. Парадигматические связи каждого дескриптора раскрываются в семантическом коде через смысловую характеристику этого дескриптора.

Наиболее известной из знаковых систем такого типа является Семантический код, разработанный в 1954—1958 гг. Дж. Перри и А. Кентом. Этот код позволяет выражать 10 разных парадигматических отношений между дескрипторами. Например, дескриптор «телефон» имеет семантический код DWCM. LQCT. MACH. TURT.001. Данный код означает, что телефон — это прибор (MACH), используемый для передачи (TURT) информации (DWCM) при помощи электричества (LQCT) [142].

Другой вариант семантического кода применен для выражения содержания понятий в ИПЯ, разработанном Э. Ф. Скороходько в Институте кибернетики АН УССР (1964 г.). Как и Семантический код Дж. Перри и А. Кента, этот код также позволяет выражать разные типы парадигматических отношений между дескрипторами (классами). Например, код $R_{001}X^0_{067}R_{007}\,(R_{001}X^0_{127}R_{011}X^0_{40})$ в данном ИПЯ означает «усилитель мощности». В этом коде буквами X и R обозначаются соответственно термины (семантические множители) и релатемы (символы отношений). В приведенном примере термины и релатемы имеют следующие фиксированные значения:

| | |
|---|---|
| $X^0_{40}$ — «мощность» | $R_{001}$ — «быть элементом класса» |
| $X^0_{67}$ — «прибор» | $R_{007}$ — «быть предназначенным» |
| $X^0_{127}$ — «увеличение» | $R_{011}$ — «иметь объект» |

Таким образом, код $R_{001}X^0_{067}R_{007}\,(R_{001}X^0_{127}R_{011}X^0_{40})$ означает, что

---

[141] M. F. Bailey, B. Lanham and J. Leibowitz. Mechanized searching in U. S. Patent Office. «Journal of the Patent Office Society», 1953, v. 35, № 7, стр. 574.

[142] J. W. Perry and A. Kent. Tools for machine literature searching. Semantic code dictionary. Equipment. Procedures. New York, Interscience Publishers, 1958, стр. 241.

усилитель мощности — это прибор ($R_{001}X^0_{67}$), предназначенный ($R_{007}$) для увеличения ($R_{001}X^0_{127}$) мощности ($R_{011}X^0_{40}$)[143].

Подводя итог рассмотрению аналитических способов выражения парадигматических отношений между дескрипторами, необходимо отметить, что родовые коды более наглядно, чем семантические, выражают родо-видовые отношения между дескрипторами или отношения вхождения (включения) между классами. Кроме того, родовые коды отличаются большей компактностью. Однако семантические коды, которые предназначаются для использования внутри машины, позволяют выражать не только родо-видовые, но и многие другие парадигматические отношения между дескрипторами.

Геометрические схемы (графы) уже давно используются для наглядного представления смысловых и иных отношений между словами. Общепринятым также является изображение перархических классификаций в виде *деревьев* — связных графов, не имеющих циклов. Однако то, что представление в виде графа словарного состава дескрипторного ИПЯ создает большие возможности для повышения эффективности информационного поиска, было в полной мере осознано лишь недавно.

В 1959 г. Л. Дойл (США) разработал концепцию документальной ИПС, основанной на активном использовании ассоциативных связей между ключевыми словами. Одним из важнейших элементов его ИПС была «ассоциативная карта» — связный граф, узлами которого являются ключевые слова, а ребрами — отрезки (стрелки), соединяющие эти ключевые слова и обозначающие ассоциативные связи (в статистическом смысле) между ними [144].

Для выражения разных степеней ассоциативной связанности ключевых слов Л. Дойл предложил использовать стрелки разного начертания и толщины, а также написание ключевых слов буквами разного размера.

На рис. 49 приведен фрагмент «ассоциативной карты», построенной Л. Дойлом на основании статистического анализа текстов 618 рефератов по психологии. Степень ассоциативной связанности каждой пары из 90 ключевых слов, встретившихся в этих рефератах, вычислялась по формуле:

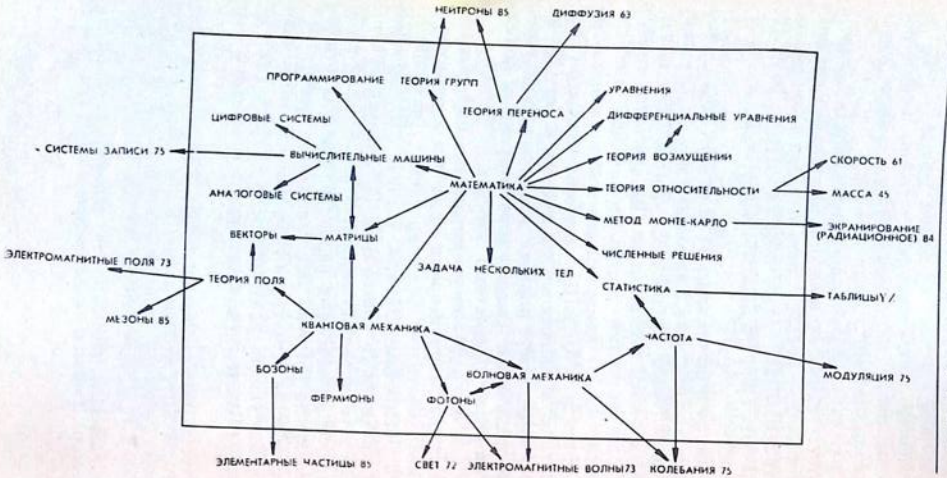
$$\rho = \frac{fN - AB}{\sqrt{AB\,(N - A)\,(N - B)}},$$

где $\rho$ — коэффициент корреляции Пирсона; $f$ — число текстов, в которых встречаются одновременно оба данных слова; $N$ —

---

[143] Т. А. Грязнухина, Л. Э. Пшеничная, Э. Ф. Скороходько. Система информационного поиска. (Об одном способе машинного поиска литературы). Киев, «Наукова думка», 1964, стр. 25.

[144] L. B. Doyle. Semantic road maps for literature searchers. «Journal of the Computing Machinery», 1961, v. 8, № 4, стр. 553—578.




| ЕВРАТОМ ЦИД ТЕЗАУРУС КЛЮЧЕВЫХ СЛОВ | МАТЕМАТИКА | | | | | | | | | | | | | | | | | ГРУППА 71 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 00 | 11 | 12 | 13 | 14 | 15 | 16 | 21 | 22 | 23 | 24 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 41 | 42 | 43 | 44 |
| 45 | 51 | 52 | 61 | 62 | 63 | 64 | 71 | 72 | 73 | 74 | 75 | 76 | 81 | 82 | 83 | 84 | 85 | 91 | 92 | 93 | 94 |

Рис. 50. Графическая схема предметного класса (поля) 71 МАТЕМАТИКА в «Тезаурусе Евратома» (изд. 1)

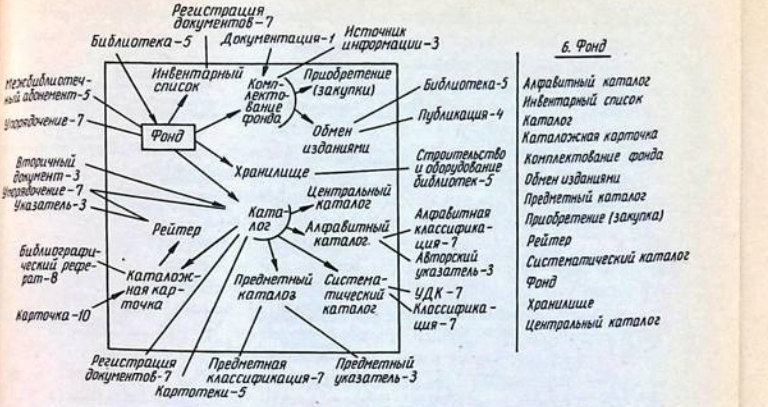



Рис. 51. Графическая схема предметного класса (поля) 6.ФОНД в «Тезаурусе для координатного индексирования [материалов] по документации и библиотековедению» (СРР, 1965 г.)

схемы. Вместо одно- и двусторонних стрелок, которые обозначали соответственно подчинительные и соподчинительные связи предметных классов, применены толстые и тонкие линии, указывающие на бóльшую или меньшую степень смысловой связанности этих классов. На схемах предметные классы изображены в виде неправильных многоугольников. Внутри каждого такого многоугольника приводится дескриптор, обозначающий данный класс, а также связанные по смыслу с этим дескриптором ключевые слова. Дескриптор дается прописными буквами, а ключевые слова — строчными. В прежние схемы «Тезауруса Евратома» (изд. 1) *ключевые слова не включались*. Наконец, из графических схем исключены все цифровые обозначения, так как они иногда ошибочно принимались за классификационные индексы.

Таким образом, в новых графических схемах «Тезауруса Евратома» (изд. 2) парадигматические связи между дескрипторами вообще никак не дифференцируются по видам (не подразделяются даже родо-видовые отношения и прочие) и отличаются друг от друга лишь по силе. Это делает такие схемы еще более похожими на «ассоциативные карты» Л. Дойла. Пример графической схемы для класса «Металлография» [150] приведен на рис. 52.

[150] Euratom-Thesaurus. Indexing terms used within Euratom's nuclear documentation system. EUR 500. e (2nd. ed., pt l). Brussels, Directorate «Dissemination of Information», Center for Information and Documentation. 1966; L. N. Rolling. A computer-aided information service for nuclear science and technology. «Journal of Documentation», 1966, v. 22, № 2, стр. 93—115.

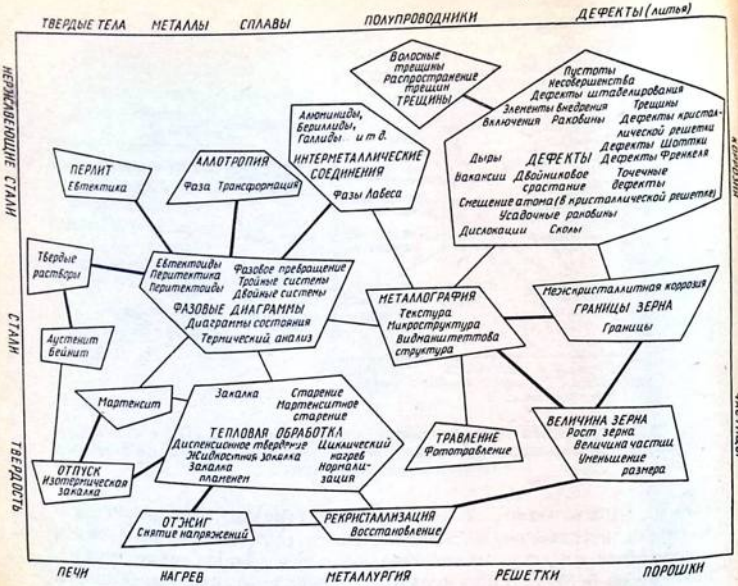



Рис. 52. Графическая схема предметного класса (поля) МЕТАЛЛОГРАФИЯ в «Тезаурусе Евратома» (изд. 2)

Другой способ графического выражения парадигматических отношений между дескрипторами применен в «Тезаурусе TDCK» (Technisch Documentatie-en Informatie-Cetrum voor Krijgsmacht — Центр технической документации и информации вооруженных сил [Нидерландов]). В четвертом издании этого тезауруса (май 1966 г.) насчитывается 11 460 дескрипторов. Путем группировки этих дескрипторов по смысловым признакам было образовано 356 непересекающихся классов [151]. Дескрипторы внутри этих классов связаны между собой родо-видовыми отношениями.

[151] J. A. Schüller, T. D. C. K. — Organisation und Systeme. Das wissenschaftliche und technische Dokumentations- und Informationszentrum der holländischen Wehrmacht. «Nachrichten für Dokumentation», 1966, 17, № 5, стр. 153—162. См. также: J. A. Schüller. The TDCK Circular Thesaurus system. «Automation and scientific communications». Washington, American Documentation Institute, 1963, Pt. 1, стр. 13—14; J. A. Schüller. De nieuwe thesaurus van het technisch documentatie-en informatie-Centrum voor krijgsmacht. «Tijdschrift voor Efficiëntie en Documentatie», 1964, v. 34; № 2, стр. 84—85.

Для графического изображения этих отношений все выделенные классы дескрипторов были представлены в виде круговых схем, на которых последовательные уровни иерархии изображены в виде концентрических окружностей. В центре окружности наименьшего радиуса помещается дескриптор, который обозначает весь данный класс дескрипторов, а на этой (или следующей) окружности размещаются дескрипторы, непосредственно подчиненные «заглавному» дескриптору, т. е. его видовые дескрипторы. Связи «заглавного» дескриптора с этими видовыми дескрипторами изображаются на схеме в виде стрелок, направленных от «заглавного» дескриптора к видовым дескрипторам.

На следующей концентрической окружности размещаются дескрипторы, которые непосредственно подчинены дескрипторам, размещенным на предыдущей окружности, и т. д. На этой же окружности могут быть размещены и некоторые дескрипторы, непосредственно подчиненные «заглавному» дескриптору.

Возможны случаи, когда дескриптор связан родо-видовым отношением с дескриптором, принадлежащим к другому классу. Тогда соответствующая родо-видовая стрелка на схеме выходит за пределы окружности наибольшего радиуса, где и приводится этот дескриптор и номер класса, к которому он принадлежит.

Составители «Тезауруса TDCK» считают, что применение на этих схемах концентрических окружностей дает примерно в пять раз больше места для размещения видовых дескрипторов, чем могут дать обычные родо-видовые деревья.

При индексировании документа исходной точкой для предметизатора служит «заглавный» дескриптор, который расположен в центре круговой схемы соответствующего класса дескрипторов. От этого дескриптора предметизатор следует в направлении родо-видовых стрелок до тех пор, пока не найдет нужный специфичный дескриптор. При этом предметизатор должен включить в поисковый образ индексируемого документа все дескрипторы, которые он встречает в данной схеме на пути следования к искомому дескриптору.

Аналогичным образом эти круговые схемы используются и при формулировании поискового предписания. Только в этом случае исходной точкой служит специфичный дескриптор, и предметизатор отыскивает его родовые дескрипторы, следуя на круговой схеме от этого специфичного дескриптора в направлении, противоположном направлению соответствующих родо-видовых стрелок.

При составлении таких круговых схем для всех выделенных классов дескрипторов в ИПЯ были полностью устранены синонимия, омонимия и полисемия слов естественного языка. Это позволило составителям «Тезауруса TDCK» выполнить его в виде собрания круговых схем, напечатанных на отдельных листах. Любая круговая схема в случае необходимости может быть заменена в «Тезаурусе TDCK» новой переработанной схемой. Выпуск

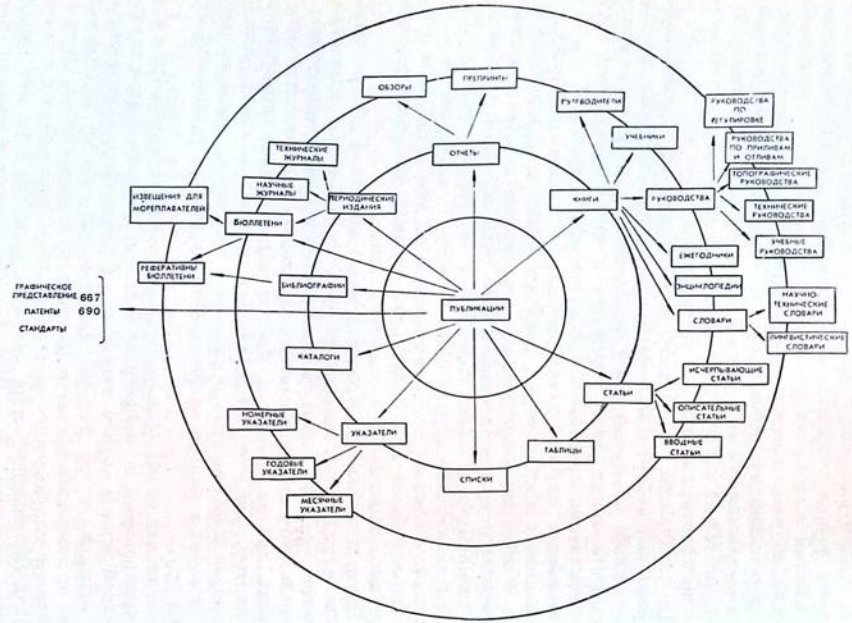



Рис. 53. Круговая схема класса публикации в «Тезаурусе TDCK»

новых круговых схем, предназначенных для замены в «Тезаурусе TDCK» устаревших схем, производится в централизованном порядке примерно четыре раза в год. Пример построения круговой схемы в «Тезаурусе TDCK» [152] приведен на рис. 53.

Итак, в тезаурусе парадигматические отношения между дескрипторами могут быть эксплицитно выражены четырьмя способами: лексикографическим, табличным, аналитическим и графическим. Наиболее наглядные из этих способов и, следовательно, наиболее удобные для потребителей информации и предметизаторов (индексаторов) — это лексикографический, табличный и графический способы, а также их разные сочетания. Однако аналитический способ является самым удобным в отношении машинных операций над дескрипторами или классами. Это означает, что если тезаурус предназначается для использования в поисковых системах, реализуемых при помощи тех или иных машин, то наилучшей формой его выполнения может быть такая, когда тезаурус состоит из двух частей:

а) справочника с наглядным отображением парадигматических отношений между дескрипторами, который предназначен для использования потребителями информации и предметизаторами (индексаторами);

б) кодовой книги с аналитическим выражением парадигматических отношений между дескрипторами, которая предназначена для преобразования поисковых образов документов и поисковых предписаний в форму, удобную для их ввода в машину.

Если для реализации ИПС используется вычислительная машина, запоминающее устройство которой имеет достаточно большую емкость и малое время обращения, то необходимость в такой кодовой книге отпадает, так как дескрипторы и обозначения парадигматических отношений, существующих между этими дескрипторами, в вычислительную машину можно вводить на естественном языке.

## ОПИСАНИЕ ОСНОВНЫХ ЗАРУБЕЖНЫХ ТЕЗАУРУСОВ

Для построения обобщенной структурной схемы тезауруса целесообразно рассмотреть основные зарубежные тезаурусы, созданные в период с 1961 по 1967 г., а именно «Тезаурус ASTIA» (изд. 2), Химико-технологический тезаурус, Тезаурус Бюро мелиорации Министерства внутренних дел США, Тезаурус технических терминов, Тезаурус AED, «Тезаурус Евратома» (изд. 2),

---

[152] L. Rolling. The role of graphic display of concept relationships in indexing and retrieval vocabularies. В кн.: P. Atherton (ed.). Classification research. Proceedings of the 2nd international study conference, (Elsinore, Denmark, 14—18th Sept. 1964). Copenhagen, Munksgaard, 1965, стр. 312.

Тезаурус Бюро кораблестроения и ремонта ВМФ США, Нормативный список предметных заголовков Американского нефтяного института, Тезаурус терминов по целлюлозе и бумаге, Тезаурус к «Engineering Index», Тезаурус по черной металлургии и Тезаурус по диффузии. При описании тезаурусов, наряду с термином «парадигматические отношения», мы будем пользоваться также более широким и менее определенным термином «ассоциативные отношения». Особое внимание обращается на выявление того, какой смысл вкладывали составители разных тезаурусов в понятие «ассоциативные дескрипторы» и какими правилами они руководствовались, когда относили те или иные дескрипторы к категории ассоциативно связанных с любым данным дескриптором. Во всех описаниях указательные пометы, применяемые для выражения в тезаурусе парадигматических отношений между дескрипторами, даются в той последовательности, в какой они используются в словарных статьях.

1. «Тезаурус ASTIA» (изд. 2) составлен для Военного документационного центра США и опубликован в декабре 1962 г. (первое издание вышло в мае 1960 г.) [153].

Этот тезаурус охватывает все отрасли науки и техники, которые нашли применение в военном деле, и содержит 5596 дескрипторов (в его первом издании насчитывалось около 6900 дескрипторов). «Тезаурус ASTIA» (изд. 2) имеет три основных раздела:

1) таблицы дескрипторов;

2) лексико-семантический указатель;

3) предметные классы (с номерным и тематическим указателями).

1.1. Таблицы дескрипторов (первый раздел тезауруса) состоят из двух частей: а) таблиц дескрипторных полей и б) таблиц дескрипторных групп. *Дескрипторное поле* — это широкий тематический класс, в который объединены соответствующие дескрипторные группы. Всего в «Тезаурусе ASTIA» 26 дескрипторных полей. В каждое дескрипторное поле входит от одной до 29 дескрипторных групп. Таблицы дескрипторных полей начинаются с пронумерованного алфавитного списка названий этих полей. Затем в таком же порядке приведены сами дескрипторные поля, состоящие из алфавитных перечней соответствующих дескрипторных групп. Названию каждой дескрипторной группы внутри поля предшествует порядковый номер этой группы. Например:

СОЦИАЛЬНЫЕ НАУКИ

44 ДОКУМЕНТАЦИЯ . . . . . 24
147 СОЦИАЛЬНЫЕ НАУКИ . . . . . 54
45 ЭКОНОМИКА . . . . . 25

---

[153] Thesaurus of ASTIA descriptors. 2nd ed. Arlington, Armed Services Technical Information Agency, 1962.

Справа указаны страницы «Тезауруса ASTIA», на которых приведены эти дескрипторные группы. Дескрипторные поля (вместе с родовыми классами, о которых уже говорилось выше) выполняют функцию путеводителя по «Тезаурусу ASTIA».

Для облегчения процедуры выбора дескрипторов, наиболее точно соответствующих смысловому содержанию документа или информационного запроса, в «Тезаурусе ASTIA» все дескрипторы сгруппированы в 170 *непересекающихся* предметных классов, называемых *дескрипторными группами*. Эти дескрипторные группы расположены в алфавитном порядке их названий и пронумерованы от 1 до 170. Внутри каждой группы дескрипторы расположены также в алфавитном порядке, причем никакой дифференциации дескрипторов по видам их парадигматических отношений с другими дескрипторами не производится. Единственным основанием для включения дескрипторов в ту или иную дескрипторную группу является их принадлежность к соответствующей предметной области. Ниже в качестве примера приводится взятая из «Тезауруса ASTIA» дескрипторная группа 44 ДОКУМЕНТАЦИЯ [154].

44 ДОКУМЕНТАЦИЯ
- АДРЕСНЫЕ КНИГИ
- БИБЛИОГРАФИИ
- ВОЕННЫЕ ПУБЛИКАЦИИ
- ГАЗЕТЫ
- ДОКУМЕНТАЦИЯ
- ЗАПИСИ
- ИНФОРМАЦИОННЫЙ ПОИСК
- КАТАЛОГИ
- МАШИННЫЙ ПЕРЕВОД
- НАСТАВЛЕНИЯ
- ОБЗОРЫ
- ОТЧЕТЫ
- ПАТЕНТЫ
- ПЕРИОДИЧЕСКИЕ ИЗДАНИЯ
- ПРЕДМЕТНОЕ ИНДЕКСИРОВАНИЕ
- РЕФЕРАТЫ
- СИМПОЗИУМЫ
- СЛОВАРИ
- СПРАВОЧНИКИ
- УКАЗАТЕЛИ
- УЧЕБНИКИ

Связи дескрипторных полей и групп со словарными статьями и родовыми классами в «Тезаурусе ASTIA» никак не обозначены.

1.2. Лексико-семантический указатель (Scope Note Index) — основной раздел «Тезауруса ASTIA». Он представляет собой словарь дескрипторов и синонимов дескрипторов, которые приведены в общем алфавитном списке. Дескрипторы в этом указателе (как и во всех остальных разделах «Тезауруса ASTIA») даны прописными

---

[154] Необходимо отметить, что дескрипторы группы 44 не соподчинены дескриптору ДОКУМЕНТАЦИЯ, так как этот дескриптор сам является членом данной группы. Это замечание справедливо и по отношению ко всем остальным дескрипторным группам.

ными буквами, а синонимы дескрипторов — строчными. Все дескрипторы и их синонимы, если это возможно, даны во множественном числе.

Под каждым дескриптором в круглых скобках приводится название дескрипторной группы, в которую входит данный дескриптор. Некоторые дескрипторы снабжены пояснительными пометами, уточняющими значение, в котором эти дескрипторы употребляются в «Тезаурусе ASTIA».

Словарные статьи многих дескрипторов содержат заголовок «Включает». Под этим заголовком даются слова и словосочетания, которые считаются синонимами «заглавного» дескриптора. Объемы понятий, выражаемых такими словами и словосочетаниями, обычно меньше или, по крайней мере, равны объемам понятий, выражаемых соответствующими дескрипторами. Приводимые под таким заглавием слова и словосочетания не могут быть дескрипторами. Поэтому они даются строчными буквами. Как правило, такие слова и словосочетания в качестве синонимов дескрипторов включены в общий алфавитный список лексико-семантического указателя.

Все синонимы дескрипторов, включенные в лексико-семантический указатель «Тезауруса ASTIA», снабжены отсылками «используй» (англ. use). Эти отсылки предписывают использовать вместо слова или словосочетания, найденного в указателе, соответствующий дескриптор (чаще всего — один), синонимом которого считается данное слово или словосочетание. Такой дескриптор либо эквивалентен по смыслу исходному слову или словосочетанию, либо близок ему по значению, либо выражает более широкое понятие. Один синоним может быть отнесен к двум и более дескрипторам, например:

Влажность атмосферы *используй* АТМОСФЕРА и ВЛАЖНОСТЬ

В словарных статьях «Тезауруса ASTIA» дескрипторы, связанные с «заглавным» дескриптором парадигматическими отношениями, группируются под заголовками «Родовые дескрипторы», «Видовые дескрипторы» и «Смотри также». Под заголовком «Смотри также» в общем алфавитном порядке приводятся дескрипторы, связанные с «заглавным» дескриптором не родо-видовыми, а другими парадигматическими отношениями. Например:

УКАЗАТЕЛИ

Смотри также:

ИНФОРМАЦИОННЫЙ ПОИСК
КАТАЛОГИ
КЛАССИФИКАЦИЯ
ПРЕДМЕТНОЕ ИНДЕКСИРОВАНИЕ

Ниже приведены два примера, из которых ясно видна схема построения словарных статей в лексико-семантическом указателе «Тезауруса ASTIA»:

1) МАГНИТНАЯ НАВИГАЦИЯ
(Навигация и управление)

*Название дескрипторной группы, в которую входит дескриптор* МАГНИТНАЯ НАВИГАЦИЯ

Родовые дескрипторы:
НАВИГАЦИЯ
Магнитная проницаемость
используй МАГНИТНЫЕ СВОЙСТВА

*Синоним дескриптора* МАГНИТНЫЕ СВОЙСТВА

МАГНИТНОЕ САМОСЖАТИЕ
(Физика плазмы)

*Название дескрипторной группы, в которую входит дескриптор* МАГНИТНОЕ САМОСЖАТИЕ

Включает:
Пинч-эффект

*Синоним дескриптора* МАГНИТНОЕ САМОСЖАТИЕ

Родовые дескрипторы:
МАГНИТОГИДРОДИНАМИКА
Смотри также:
ФИЗИКА ПЛАЗМЫ
МАГНИТНЫЕ СВОЙСТВА
(Электричество и магнетизм)

*Другие ассоциативные дескрипторы*

*Название дескрипторной группы, в которую входит дескриптор* МАГНИТНЫЕ СВОЙСТВА

(Реакция оборудования или материалов на магнитные поля, а также способность оборудования или материалов создавать магнитные поля)

*Уточнение значения, в котором дескриптор* МАГНИТНЫЕ СВОЙСТВА *употребляется в тезаурусе*

Включает:
Диамагнитные свойства
Магнитная восприимчивость
Магнитная проницаемость

*Синонимы дескриптора* МАГНИТНЫЕ СВОЙСТВА

Видовые дескрипторы:
МАГНИТОСТРИКЦИЯ
ПРОНИЦАЕМОСТЬ МАГНИТНАЯ
ЭФФЕКТ ХОЛЛА
Смотри также:
ГИСТЕРЕЗИС
ИНДУКЦИЯ
МАГНЕТИЗМ
ПАРАМАГНИТНЫЙ РЕЗОНАНС
ПОЛЯРИЗАЦИЯ

*Другие ассоциативные дескрипторы*

2) УГЛЕРОД
(Химические элементы)

*Название дескрипторной группы, в которую входит дескриптор* УГЛЕРОД

Включает:
Активированный уголь — *Синоним дескриптора* УГЛЕРОД
Родовые дескрипторы:
НЕМЕТАЛЛЫ
ЭЛЕМЕНТЫ IV ГРУППЫ
Видовые дескрипторы:
АЛМАЗЫ
ГАЗОВАЯ САЖА
ГРАФИТ
Смотри также:
ДРЕВЕСНЫЙ УГОЛЬ — *Другие ассоциативные дескрипторы*
КАРБИДЫ

Необходимо отметить, что при использовании «Тезауруса ASTIA» названия сплавов составляются из дескрипторов типа N, СПЛАВЫ (N — название металла, входящего в сплав). Например, если речь идет о сплаве меди с оловом, то нужно использовать дескрипторы МЕДЬ, СПЛАВЫ+ОЛОВО, СПЛАВЫ. В «Тезаурусе ASTIA» этим дескрипторам соответствуют английские словосочетания Copper Alloys и Tin Alloys.

Вместо названий неорганических соединений, в «Тезаурусе ASTIA» даны дескрипторы двух типов. Эти дескрипторы обозначают радикалы или ионы, образующие соответственно положительные и отрицательные части молекул. Дескрипторы первого типа — это словосочетания, в которых после названия элемента или группы элементов стоит слово «соединения» или «радикалы» (имеются в виду английские словосочетания типа Halogen Compounds, Uranyl Radicals и т. д.). В качестве дескрипторов второго типа используются общепринятые названия анионов. При помощи дескрипторов этих двух типов и выражаются названия неорганических соединений. Например, если в документе или информационном запросе говорится о гидроокиси алюминия, то необходимо использовать дескрипторы АЛЮМИНИЙ, СОЕДИНЕНИЯ+ГИДРООКИСИ, если же речь идет о хлорплутонате цезия, то используются дескрипторы ЦЕЗИЙ, СОЕДИНЕНИЯ+ПЛУТОНАТЫ+ХЛОРИДЫ, и т. д.

Дескрипторы, обозначающие органические соединения или предназначенные для построения их названий на ИПЯ, а также соответствующие правила даны в специальном приложении к «Тезаурусу ASTIA», которое названо «Химическим тезаурусом ASTIA» (содержит 1,6 тыс. дескрипторов). В качестве дескрипторов в нем используются лишь названия наиболее важных органических соединений — лекарственных и биологических препаратов, энзимов, витаминов и т. п. Названия остальных органических соединений составляются из дескрипторов, обозначающих фрагменты структурных формул этих соединений, т. е. их функциональные или замещенные группы, радикалы и ионы.

1.3. Предметные классы, имеющие форму таблиц, построены путем обобщения родо-видовых отношений, которые существуют между отдельными дескрипторами и показаны в словарных статьях лексико-семантического указателя. Схема построения предметных классов, используемых в «Тезаурусе ASTIA», была достаточно подробно описана выше. Поэтому в данном разделе можно ограничиться лишь описанием указателей к этим предметным классам.

*Номерной указатель предметных классов.* В этом указателе номерам предметных классов поставлены в соответствие названия этих классов.

*Тематический указатель к предметным классам.* В этом указателе под широкими тематическими заголовками (всего таких заголовков 13) даны в алфавитном порядке названия соответствующих предметных классов с их порядковыми номерами.

Например:

| | |
|---|---|
| ОБЩАЯ И ФИЗИЧЕСКАЯ ХИМИЯ | |
| ГЕОГРАФИЯ И ГЕОЛОГИЯ | 49 |
| МАТЕРИАЛЫ | 37 |
| МЕТАЛЛУРГИЯ | 69 |
| НЕОРГАНИЧЕСКАЯ ХИМИЯ | 64 |
| ПРОЦЕССЫ И ОПЕРАЦИИ | 72 |
| ТЕПЛОТА И ТЕРМОДИНАМИКА | 36 |
| ТОПЛИВО И ГОРЕНИЕ | 15 |
| ФИЗИЧЕСКИЕ СВОЙСТВА | 32 |
| ФОТОГРАФИЯ | 43 |
| ХИМИЧЕСКИЕ ЭЛЕМЕНТЫ | 65 |
| ХИМИЯ И ХИМИЧЕСКИЕ РЕАКЦИИ | 66 |

Порядок расположения заголовков в тематическом указателе — алфавитный.

2. «Химико-технологический тезаурус» — официальное издание Американского института инженеров-химиков [155] — вышел из печати в июне 1961 г. В предисловии к «Химико-технологическому тезаурусу» говорится, что этот тезаурус представляет собой «средство для напоминания слов или список ассоциированных слов (word association list), предназначенный для повышения качества индексирования и улучшения поиска информации. . . Основное назначение этого тезауруса — снабжать индексаторов дополнительными терминами, которые могут быть использованы ими для индексирования технических документов. Этот тезаурус отражает различные аспекты (viewpoints), родо-видовые связи и семантику терминов, выявленные путем рассмотрения каждого термина редакционным

[155] Chemical engineering thesaurus. New York, American Institute of Chemical Engineers, 1961.

комитетом»[156]. В «Химико-технологическом тезаурусе» содержится более 7 тыс. дескрипторов.

Тезаурус состоит из двух частей: Часть I. Химико-технологические термины. Часть II. Химические термины. Первую часть тезауруса составляют употребляемые в химической литературе дескрипторы общего характера, а вторую — названия химических элементов и соединений, сплавов, материалов и т. д.

В «Химико-технологическом тезаурусе» словарная статья дескриптора имеет следующую структуру. Сначала под «заглавным» дескриптором приводятся все слова, считающиеся в тезаурусе синонимами этого дескриптора. Каждое такое слово отмечено в словарной статье аббревиатурой SF (*англ.* seen from= смотри от).

Затем дается список всех дескрипторов, которые в родо-видовом дереве стоят ниже (в том числе непосредственно ниже) «заглавного» дескриптора. Видовые дескрипторы в словарной статье отмечены аббревиатурами GT (англ. generic to=родовой по отношению к).

Следующая группа — дескрипторы, которые в родо-видовом дереве стоят выше (в том числе и непосредственно выше) «заглавного» дескриптора. Родовые дескрипторы в словарной статье отмечены аббревиатурами PO (англ. post on=приписать к [«заглавному» дескриптору]).

В конце словарной статьи в общем алфавитном порядке приводятся дескрипторы, связанные с «заглавным» дескриптором не родо-видовыми, а другими ассоциативными отношениями. Каждый из этих дескрипторов отмечен аббревиатурой RT (англ. related terms=родственные термины). В связи с этим необходимо отметить, что в «Химико-технологическом тезаурусе» родственными (related) с «заглавным» дескриптором считаются такие дескрипторы, которые не являются синонимами «заглавного» дескриптора, не связаны с ним отношениями «род—вид» или «вид—род», но близки «заглавному» дескриптору по значению или часто ассоциируются с ним[157]. Ниже дан пример группировки в словарной статье таких родственных дескрипторов.

УКАЗАТЕЛИ

| | | | |
|---|---|---|---|
| RT | ДЕСКРИПТОРЫ | RT | КАРТОЧНЫЕ РЕЙТЕРЫ |
| RT | ДИАМИКРОКАРТЫ ТИПА «МИНИКАРД» | | ТИПА «РАДЕКС» |
| | | RT | КАРТЫ БАТТЕНА |
| RT | ИНФОРМАЦИЯ | RT | КЛАССИФИКАЦИЯ |
| RT | КАТАЛОГИ | RT | ПЕРФОКАРТЫ |
| RT | КАРТОТЕКИ | RT | СПИСКИ |
| RT | КАРТОЧКИ | RT | «УНИТЕРМ» |

[156] Chemical engineering thesaurus. New York, American Institute of Chemical Engineers, 1961, стр. 1.
[157] Там же, стр. 11.

Порядок расположения дескрипторов и их синонимов в перечисленных подразделах словарной статьи — алфавитный. Любой из этих подразделов может быть пустым (т. е. не содержать ни одного элемента).

Названия неорганических соединений в «Химико-технологическом тезаурусе» приводятся в алфавитном порядке:

а) под названиями их катионообразующих элементов;

б) под названиями их родовых групп, которые обозначаются такими общепринятыми названиями анионов, как *хлориды*, *сульфаты*, *оксиды* и т. д.

Для органических соединений в этом тезаурусе используются названия, которые наиболее широко применяются химиками. Однако эти названия, как правило, соответствуют Женевской номенклатуре. Органические соединения приводятся также под названиями их родовых классов (например, под такими, как *амиды*, *амины*, *спирты*, *эфиры простые*, *эфиры сложные* и т. д.).

Другие названия, употребляемые для обозначения различных химических соединений, считаются синонимами основных названий и связаны с ними отсылкой «смотри» (англ. see).

При расстановке дескрипторов и их синонимов в алфавитном порядке не учитываются интервалы между словами в словосочетаниях, арабские цифры, пунктуационные знаки, химические символы, греческие и латинские аффиксы (когда они приводятся не полностью, а в виде соответствующих букв латинского алфавита), а также словосочетания, начинающиеся с отсылки «смотри». Все пояснительные пометы, уточняющие значения отдельных дескрипторов, включаются в общий алфавитный порядок.

В начале каждого буквенного раздела тезауруса сгруппированы (в углубленном алфавитном порядке) соответствующие аббревиатуры, буквенные и буквенно-цифровые обозначения, а также все термины, начинающиеся с однобуквенного сокращения или условного обозначения (например, G-VALUE, J-FACTORS, O-RINGS и т. д.).

Без сокращений в «Химико-технологическом тезаурусе» даны:

а) названия стран и штатов, когда эти названия выступают в качестве терминов;

б) названия стран, когда эти названия идентифицируют или локализуют остальную часть словосочетания (названия штатов США в этом случае сокращаются);

в) названия компаний и корпораций.

В «Химико-технологическом тезаурусе» все дескрипторы и их синонимы, обозначающие перечислимые предметы и сочетающиеся с количественными числительными, даны во множественном числе.

При составлении этого тезауруса группировка терминов в дескрипторные статьи производилась при помощи ЭЦВМ. Рукопись тезауруса, напечатанная выводным устройством этой

ЭЦВМ, была использована в качестве оригинал-макета для фотоофсетного размножения. Поэтому весь «Химико-технологический тезаурус» набран одноразмерным шрифтом, т. е. только прописными буквами. Ниже приводятся примеры построения словарных статей в «Химико-технологическом тезаурусе».

1) ПУБЛИКАЦИИ
SF БРОШЮРЫ
SF БУКЛЕТЫ
SF ПЕРИОДИЧЕСКИЕ ИЗДАНИЯ
GT РУКОВОДСТВА [ПУБЛИКАЦИИ]
PO СООБЩЕНИЯ
RT ИНФОРМАЦИЯ
RT КАТАЛОГИ
RT ЛИТЕРАТУРА
RT ОЗНАКОМИТЕЛЬНАЯ РАССЫЛКА ПУБЛИКАЦИЙ
RT ОТЧЕТЫ
RT ОФИЦИАЛЬНЫЕ СООБЩЕНИЯ ДЛЯ ПЕЧАТИ
RT ПРОТОКОЛЫ
RT ЦИРКУЛЯРЫ

2) ЖИРНЫЕ КИСЛОТЫ
SF ГЛИЦЕРИДЫ
GT ОЛЕИНОВЫЕ КИСЛОТЫ
GT ПАЛЬМИТИНОВЫЕ КИСЛОТЫ
GT РИЦИНОЛЕИНОВЫЕ КИСЛОТЫ
GT СТЕАРИНОВЫЕ КИСЛОТЫ
PO КИСЛОТЫ [КАРБОНОВЫЕ]

3) ОКИСЬ ДЕЙТЕРИЯ
SF ТЯЖЕЛАЯ ВОДА
PO ОКИСЛЫ [НЕОРГАНИЧЕСКИЕ]
RT ВОДА
RT ВОДОРОД
RT ИЗОТОПЫ

4) БОРАН
БОРАТЫ [ЭФИРЫ СЛОЖНЫЕ]
PO ЭФИРЫ СЛОЖНЫЕ
БОРАТЫ [СОЛИ]
ПЕРБОРАТЫ
GT БОРНОКИСЛЫЙ НАТРИЙ
GT ОКСИФТОРОБОРАТ НАТРИЯ
GT ПЕРБОРНОКИСЛЫЙ БАРИЙ
GT ПЕРБОРНОКИСЛЫЙ НАТРИЙ
БОРНАЯ КИСЛОТА
SF САССОЛИТ
PO КИСЛОТЫ [НЕОРГАНИЧЕСКИЕ]
БОРНОКИСЛЫЙ НАТРИЙ

SF БУРА
SF ТЕТРАБОРАТ НАТРИЯ
PO БОРАТЫ [СОЛИ]
БОРНЫЙ АНГИДРИД — СМОТРИ ОКИСЬ БОРА
БУРА — СМОТРИ БОРНОКИСЛЫЙ НАТРИЙ

3. Тезаурус Бюро мелиорации Министерства внутренних дел США (предварительное издание) [158] опубликован в октябре 1963 г. В нем насчитывается 3279 дескрипторов и 3200 ключевых слов. Тезаурус состоит из следующих частей: Часть 1. Дескрипторы и их синонимы. Часть 2. Дескрипторные группы. Часть 3. Дескрипторные поля.

3. 1. Дескрипторы и их синонимы даны в одном общем списке. Порядок их расположения в этом списке — алфавитный. Синонимы выделены путем их сдвига вправо на два знака по отношению к дескрипторам. Ниже для иллюстрации приведен фрагмент из этого алфавитного списка дескрипторов и синонимов:

ЭЛЕКТРОМАГНЕТИЗМ GP 36, СМОТРИ ТАКЖЕ GP 13
ЭЛЕКТРОМАГНИТНОЕ ЭКРАНИРОВАНИЕ GP 13
ЭЛЕКТРОМАГНИТНЫЕ ВОЛНЫ УПОТРЕБЛЯЙ ЭЛЕКТРОМАГНЕТИЗМ
ЭЛЕКТРОМАГНИТНЫЕ ИМПУЛЬСЫ GP 13
ЭЛЕКТРОМАГНИТНЫЕ СВОЙСТВА GP 37, СМОТРИ ТАКЖЕ GP 13
ЭЛЕКТРОМАГНИТЫ GP 13, СМОТРИ ТАКЖЕ GP 36
ЭЛЕКТРОННАЯ МИКРОСКОПИЯ GP 13
ЭЛЕКТРОННЫЕ ВЫЧИСЛИТЕЛЬНЫЕ МАШИНЫ УПОТРЕБЛЯЙ ВЫЧИСЛИТЕЛЬНЫЕ МАШИНЫ . . . . . .
ЭЛЕКТРОННЫЕ ЛАМПЫ GP 13
ЭЛЕКТРОННЫЕ МИКРОСКОПЫ УПОТРЕБЛЯЙ ЭЛЕКТРОННАЯ МИКРОСКОПИЯ . . . . . .
ЭЛЕКТРОННЫЕ ПУЧКИ GP 33
ЭНЕРГИЯ GP 36
ЭНЕРГИЯ [АТОМНАЯ] УПОТРЕБЛЯЙ ЯДЕРНАЯ ЭНЕРГИЯ
ЭНЕРГИЯ [КИНЕТИЧЕСКАЯ] УПОТРЕБЛЯЙ КИНЕТИЧЕСКАЯ ЭНЕРГИЯ . . . . . .
ЭНЕРГИЯ [ПРИЛИВНАЯ] УПОТРЕБЛЯЙ ПРИЛИВНАЯ ЭНЕРГИЯ
ЭНЕРГИЯ [СОЛНЕЧНАЯ] УПОТРЕБЛЯЙ СОЛНЕЧНАЯ РАДИАЦИЯ
ЭНЕРГИЯ [ЯДЕРНАЯ] УПОТРЕБЛЯЙ ЯДЕРНАЯ ЭНЕРГИЯ

---

[158] Thesaurus of descriptors. A list of keywords and crossreferences for indexing and retrieving the literature of water resources development. Tentative ed., Denver, Colo., U. S. Department of the Interior, Bureau of Reclamation, 1963; B. P. Belport. Bureau of Reclamation thesaurus of descriptors. «Bull. de l'Association internationale des documentalistes et techniciens de l'information», 1966, v. 5, № 4, стр. 53—57. Далее название этого издания будем давать в следующем виде: «Bull. de l'AID».

3. 2. Отличительной особенностью Тезауруса Б
ции является то, что в нем родо-видовые и прочие
дескрипторы в словарных статьях (т. е. под «заглав
тором) не приводятся. Вместо этого каждый дескрип
части, как это видно из приведенных выше приме
ссылкой на номер дескрипторной группы, членом
является. В тезаурусе это имеет следующий вид:

ИНФОРМАЦИОННЫЙ ПОИСК GP 10

Аббревиатура GP здесь означает слово «гру
GP=group). Ниже полностью приведена взятая и
Бюро мелиорации дескрипторная группа 10 ДОКУМ
в которую входит дескриптор ИНФОРМАЦИОННЫ

| GP 10 | ОПРОСНЫЕ ЛИСТЫ |
|---|---|
| ДОКУМЕНТАЦИЯ | ПАМЯТЬ |
| БИБЛИОГРАФИИ | ПЕРЕВОД [ЯЗЫКОВЫ |
| БИБЛИОТЕКИ | ПЕРФОЛЕНТЫ |
| БИБЛИОТЕЧНЫЕ КЛАССИФИКАЦИИ | ПЕЧАТАНИЕ |
| | ПИСАНИЕ НАУЧНО-Т |
| ГРУППИРОВКА ПО ЗНАЧЕНИЮ | СКИХ РАБОТ |
| ДОКУМЕНТАЦИЯ | ПРЕДМЕТНОЕ ИНДЕ |
| ЗАПИСИ | НИЕ |
| ИНФОРМАЦИОННЫЙ ПОИСК | ПРОФИЛИ ИНТЕРЕСОВ |
| КАТАЛОГИ | РЕДАКТИРОВАНИЕ |
| КОДИРОВАНИЕ | РЕПРОДУЦИРОВАНИЕ |
| КОДЫ | ТАНИЕ] |
| ЛЕКСИКА | РЕФЕРАТЫ |
| МАГНИТНАЯ ЗАПИСЬ | СЕМАНТИКА |
| МАГНИТНЫЕ ЛЕНТЫ | СЛОВАРИ |
| МАШИННЫЙ ПЕРЕВОД | СПРАВОЧНИКИ |
| МИКРОФИЛЬМЫ | УКАЗАТЕЛИ |
| НОМЕНКЛАТУРЫ | ШУМ [ПОИСКОВЫЙ] |
| ОБЗОРЫ | ЯЗЫКИ |

Необходимо отметить, что некоторые дескрипторы в
ной части Тезауруса Бюро мелиорации имеют дополнит
ссылки «Смотри также» (see also) на номера групп, в котор
непосредственно не входят, но с которыми имеют опреде
смысловые связи. Например:

ИСПЫТАНИЕ МАТЕРИАЛОВ GP 31, СМОТРИ ТАКЖЕ GP 44 (G
МЕХАНИЧЕСКИЕ СВОЙСТВА, GP 44 = ПРОЧНОСТЬ МАТЕРИ

Всего в Тезаурусе Бюро мелиорации образовано 47 дес
торных групп. В предисловии к этому тезаурусу подчерк
что объединение дескрипторов в предметно-тематические гр

производилось не на основе их родо-видовых отношений, а на основе «вероятности их ассоциаций в индексируемой литературе» [159].

Дескрипторные группы даны во второй части тезауруса. Сначала приводится пронумерованный алфавитный список названий дескрипторных групп с индексами дескрипторных полей, в которые входят эти группы.

Например:

29. МАТЕМАТИКА И ВЫЧИСЛИТЕЛЬНЫЕ МАШИНЫ (E)
28. МАТЕРИАЛЫ (D)
30. МАШИНОСТРОЕНИЕ (C)

и т. д.[160]

Затем в таком же порядке приводятся сами дескрипторные группы.

3. 3. В тезаурусе образовано также 6 дескрипторных полей, в каждое из которых входит от 5 до 10 дескрипторных групп. Ниже приводится список этих дескрипторных полей.

A. ГРАЖДАНСКОЕ СТРОИТЕЛЬСТВО
B. НАУКИ О ЗЕМЛЕ И СЕЛЬСКОЕ ХОЗЯЙСТВО
C. ЭЛЕКТРОТЕХНИКА, МАШИНОСТРОЕНИЕ И ПРОМЫШЛЕННАЯ ТЕХНИКА
D. МАТЕРИАЛОВЕДЕНИЕ
E. ФИЗИЧЕСКИЕ, ХИМИЧЕСКИЕ И БИОЛОГИЧЕСКИЕ НАУКИ
F. СОЦИАЛЬНЫЕ НАУКИ И ПРОЧИЕ ВОПРОСЫ

Дескрипторная группа 10 ДОКУМЕНТАЦИЯ входит в дескрипторное поле F. СОЦИАЛЬНЫЕ НАУКИ И ПРОЧИЕ ВОПРОСЫ. В предисловии к Тезаурусу Бюро мелиорации говорится, что «эти шесть дескрипторных полей с входящими в них группами ... дают тому, кто пользуется данным тезаурусом, общее представление о предметной широте тезауруса. Но особую ценность они должны представлять для того, кто производит информационный поиск: при описании предметного содержания искомого материала эти поля помогают ему выбирать такие же дескрипторы, какие, по всей вероятности, использовались и предметизатором» [161].

Третья часть Тезауруса Бюро мелиорации начинается с алфавитного списка названий дескрипторных полей, после чего в такой же последовательности даны сами дескрипторные поля. Порядок расположения дескрипторных групп в каждом дескрипторном поле — алфавитный. Например, дескрипторное поле E.

[159] Thesaurus of descriptors. . ., стр. I.

[160] В приводимых примерах номера дескрипторных групп и буквенные обозначения дескрипторных полей соответствуют алфавитному порядку названий этих групп и полей на английском языке.

[161] Там же, стр. II.

в Тезаурусе Бюро мелиорации имеет следующий вид:

Е. ФИЗИЧЕСКИЕ, ХИМИЧЕСКИЕ И БИОЛОГИЧЕСКИЕ НАУКИ


| Номер группы | Группа | Стр. |
|---|---|---|
| 04. | Биология и медицина | 107 |
| 29. | Математика и вычислительные машины | 118 |
| 36. | Физические и физико-химические понятия | 120 |
| 05. | Химия и химическая технология | 107 |
| 33. | Ядерная и атомная физика | 120 |


(Справа указаны страницы, на которых в Тезаурусе Бюро мелиорации даны названные дескрипторные группы).

4. «Тезаурус технических терминов» опубликован Объединенным советом инженеров США в июне 1964 г. В предисловии к этому тезаурусу говорится, что «в противоположность словарю тезаурус не раскрывает точного значения слов; показывая каждый термин в его связи с другими терминами, он служит как бы «путеводителем» по технической терминологии» [162]. «Тезаурус технических терминов», в который включено 10 515 дескрипторов и ключевых слов (дескрипторов — около 7750), имеет такую же внутреннюю структуру, как и описанный ранее «Химико-технологический тезаурус». Порядок расположения дескрипторов и их синонимов в тезаурусе — алфавитный. Синонимы связаны с подставляемыми вместо них дескрипторами при помощи отсылки «используй» (*англ.* use). В словарных статьях «Тезауруса технических терминов» для обозначения синонимичных ключевых слов, а также видовых, родовых и прочих ассоциативных дескрипторов используются следующие аббревиатуры:

1) UF (*англ.* used for=используется вместо. . .) — для обозначения ключевых слов, которые в данном тезаурусе считаются синонимами соответствующего «заглавного» дескриптора;

2) NT (*англ.* narrower term=более узкий [по значению] термин) — для обозначения видовых дескрипторов;

3) BT (*англ.* broader term=более широкий [по значению] термин) — для обозначения родовых дескрипторов;

4) RT (*англ.* related term=родственный термин) — для обозначения прочих ассоциативных дескрипторов.

Особого рассмотрения заслуживают смысловые отношения, обозначаемые в этом тезаурусе пометой RT (родственный термин). В правилах составления «Тезауруса технических терминов» этот тип отношений между дескрипторами интерпретируется следующим образом: «Помета РОДСТВЕННЫЙ ТЕРМИН используется для обозначения такой связи между дескрипторами, которая

[162] Thesaurus of engineering terms: a list of engineering terms and their relationships for use in vocabulary control, in indexing and retrieving engineering information, 1st ed. New York, Engineers Joint Council, 1964.

не поддается четкому определению (a non-structured relationship)... Вообще два любых дескриптора связываются отношением RT, когда имеется уверенность в том, что потребитель при рассмотрении одного термина может испытывать желание, чтобы ему напомнили о существовании и другого дескриптора» [163].

В «Тезаурусе технических терминов» помета «родственный термин» используется для обозначения:

а) дескрипторов, которые близко связаны по смыслу;

б) дескрипторов, которые являются полусинонимами;

в) дескрипторов, между которыми существует отношение «точки зрения» (например, отношение широкого дескриптора к более узким, которое основано на использовании для разных целей предмета, обозначенного широким дескриптором). Отношение такого типа существует между дескриптором СПИРТЫ и дескрипторами АНТИФРИЗЫ, РАСТВОРИТЕЛИ и т. д.;

г) дескрипторов, связанных отношением «часть—целое» («целое—часть»);

д) дескрипторов, которые являются членами разных иерархических структур, но по смыслу связаны друг с другом.

Помета «родственный термин» не применяется для обозначения связей между дескрипторами, стоящими на разных уровнях одной и той же иерархии. Например:

УМЕНЬШЕНИЕ или УМЕНЬШЕНИЕ
RT КОНТРОЛЬ RT КОНТРОЛЬ ЗА ЗАГРЯЗНЕНИЯМИ
RT РЕГУЛИРОВАНИЕ ПАВОДКОВ

но не

УМЕНЬШЕНИЕ
RT КОНТРОЛЬ
RT КОНТРОЛЬ ЗА ЗАГРЯЗНЕНИЯМИ
RT РЕГУЛИРОВАНИЕ ПАВОДКОВ

В приведенном примере пометой RT могут быть связаны с дескриптором УМЕНЬШЕНИЕ либо дескриптор КОНТРОЛЬ, либо непосредственно подчиненные ему дескрипторы КОНТРОЛЬ ЗА ЗАГРЯЗНЕНИЯМИ и РЕГУЛИРОВАНИЕ ПАВОДКОВ, но не дескрипторы обоих уровней этой иерархии. В данной альтернативе первое предпочтительно тогда, когда многие видовые дескрипторы являются «родственными» (в указанном выше смысле) данному дескриптору.

Все аббревиатуры RT обозначают симметричные отношения, за исключением случаев, когда дескрипторы имеют ограничительную помету «Данный термин вообще не следует использовать — употребляй родственный термин». Дескриптор с такой ограничительной пометой может быть связан отношением RT

[163] E. Y. Speight. What is «The Thesaurus of engineering terms» developed by Engineers' Joint Council (EJC) «Bull. de l'AID», 1966, v. 5, № 4, стр. 42.

с другим дескриптором, если последний снабжен такой же пометой [164]. Ниже приведены примеры построения словарных статей в «Тезаурусе технических терминов».

1) АТОМНАЯ ЭНЕРГИЯ
УПОТРЕБЛЯЙ ЯДЕРНАЯ ЭНЕРГИЯ
АТОМНЫЕ РЕАКТОРЫ
УПОТРЕБЛЯЙ ЯДЕРНЫЕ РЕАКТОРЫ
АТОМНЫЕ СВОЙСТВА
UF АТОМНЫЙ ВЕС
АТОМНЫЙ НОМЕР
МАССА АТОМА
МАССОВОЕ ЧИСЛО
NT СТРОЕНИЕ АТОМА
RT АТОМЫ
КВАНТОВАЯ ТЕОРИЯ
КВАНТОВЫЕ ЧИСЛА
ПОТЕНЦИАЛЫ ИОНИЗАЦИИ
СВЕРХТОНКАЯ СТРУКТУРА
ЯДЕРНАЯ ФИЗИКА
АТОМНЫЙ НОМЕР
УПОТРЕБЛЯЙ АТОМНЫЕ СВОЙСТВА

2) ВЫЧИСЛИТЕЛЬНЫЕ МАШИНЫ
UF МЕТОДЫ ПРИМЕНЕНИЯ ВЫЧИСЛИТЕЛЬНЫХ МАШИН
СПЕЦИАЛИЗИРОВАННЫЕ ВЫЧИСЛИТЕЛЬНЫЕ МАШИНЫ
УНИВЕРСАЛЬНЫЕ ВЫЧИСЛИТЕЛЬНЫЕ МАШИНЫ
NT АНАЛОГОВЫЕ ВЫЧИСЛИТЕЛЬНЫЕ МАШИНЫ
АНАЛОГО-ЦИФРОВЫЕ ВЫЧИСЛИТЕЛЬНЫЕ МАШИНЫ
ЦИФРОВЫЕ ВЫЧИСЛИТЕЛЬНЫЕ МАШИНЫ
BT МАТЕМАТИЧЕСКИЕ МАШИНЫ &
RT АВТОМАТИЗАЦИЯ
ДЕТАЛИ ВЫЧИСЛИТЕЛЬНЫХ МАШИН &
ИСКУССТВЕННЫЙ РАЗУМ
КИБЕРНЕТИКА
КОДИРОВАНИЕ
ЛОГИЧЕСКИЕ СХЕМЫ
ЛОГИЧЕСКИЙ СИНТЕЗ
МАГНИТНАЯ ЛЕНТА
МАШИННЫЙ ПЕРЕВОД
ОБРАБОТКА ДАННЫХ
ПРЕОБРАЗОВАТЕЛИ ДАННЫХ
ПРОГРАММЫ (ДЛЯ ВЫЧИСЛИТЕЛЬНЫХ МАШИН)

[164] E. Y. Speight. What is «The Thesaurus of engineering terms» developed by Engineers'. Joint Council (EJC) «Bull. de l'AID», 1966, v. 5, № 4, стр. 42.

СИСТЕМЫ (ДЛЯ) ОБРАБОТКИ ДАННЫХ
СЧЕТНЫЕ ЛИНЕЙКИ
СЧЕТНЫЕ МАШИНЫ
ТЕОРИЯ АВТОМАТОВ
ТЕОРИЯ ИНФОРМАЦИИ

3) МАШИННЫЙ ПЕРЕВОД
UF МЕХАНИЧЕСКИЙ ПЕРЕВОД
ПЕРЕВОДЧЕСКИЕ МАШИНЫ
BT ПЕРЕВОД
RT ВЫЧИСЛИТЕЛЬНЫЕ МАШИНЫ
ЛИНГВИСТИКА
ТЕОРИЯ ИНФОРМАЦИИ

4) ОПТИЧЕСКИЕ СВОЙСТВА
(ВКЛЮЧАЕТ СВОЙСТВА ВИДИМЫХ, ИНФРАКРАСНЫХ И УЛЬТРАФИОЛЕТОВЫХ ЭЛЕКТРОМАГНИТНЫХ ВОЛН И ИХ ЭФФЕКТЫ)
UF ОПТИЧЕСКАЯ ПРОЗРАЧНОСТЬ
BT ФИЗИЧЕСКИЕ СВОЙСТВА
RT АЛЬБЕДО
ВИДИМОСТЬ
ВРАЩЕНИЕ ПЛОСКОСТИ ПОЛЯРИЗАЦИИ
ДВОЙНОЕ ЛУЧЕПРЕЛОМЛЕНИЕ
ДИСПЕРСИЯ (ВОЛН)
ДИФРАКЦИЯ
ДИХРОИЗМ
ИНТЕРФЕРЕНЦИЯ
МУТНОСТЬ
НЕПРОЗРАЧНОСТЬ
ОПТИЧЕСКИЕ ИЗМЕРЕНИЯ
ОТРАЖАТЕЛЬНАЯ СПОСОБНОСТЬ
ПОЛЯРИЗАЦИЯ
ПРЕЛОМЛЯЮЩАЯ СПОСОБНОСТЬ
ПРОЗРАЧНОСТЬ
СВЕТ (ОСВЕЩЕНИЕ)
СВЕТИМОСТЬ
СТЕКЛО &
ФОТОГАЛЬВАНИЧЕСКИЙ ЭФФЕКТ
ФОТОПРОВОДИМОСТЬ
ФОТОЭЛЕКТРИЧЕСКИЙ ЭФФЕКТ
ФОТОЭМИССИОННАЯ СПОСОБНОСТЬ
ЭЛЕКТРОМАГНИТНОЕ ПОГЛОЩЕНИЕ
ЯРКОСТЬ

5) ОПТИЧЕСКИЕ ПИРОМЕТРЫ #
BT ИЗМЕРИТЕЛЬНЫЕ ПРИБОРЫ &
ОПТИЧЕСКИЕ ПРИБОРЫ &

ПРИБОРЫ ДЛЯ ИЗМЕРЕНИЯ
RT ОПТИЧЕСКИЕ ИЗМЕРЕНИЯ
ПИРОМЕТРЫ ИЗЛУЧЕНИЯ
ЦВЕТОВАЯ ТЕМПЕРАТУРА

Из приведенных примеров видно, что
ческих терминов» некоторые дескрипторы
# и &. Символ # указывает на то, что
риптор может быть заменен одним или не
дескрипторами, которые выражают более ши
мечены символом &. Наоборот, символом &
рипторы, которые можно использовать вмест
скольких других дескрипторов, выражающих бол

5. «Тезаурус AED» [165] был опублик
1963 г. Центром документации по атомной энер
туте им. Гмелина (ФРГ). «Тезаурус AED» сост
щих трех частей: [166]

Часть 1. *Материалы и химические вещества* (
звания: химических элементов и их групп, химич
нений и их групп, сплавов и материалов). Часть
*и понятия общего характера* (включает названия: с
ментных частиц, электромагнитных излучений, химич
цессов, приборов и аппаратов). Часть 3. *Названия, о*
*и типы реакторов.*

В «Тезаурусе AED» все дескрипторы, обозначающие
лимые предметы, даются во множественном числе. В
части тезауруса дескрипторы и их синонимы приведены в
общем списке. Дескрипторы даны в этом списке прописным
вами, а синонимы дескрипторов — строчными. Порядок по
ния списков — алфавитный. Этот порядок нарушается тол
двух следующих случаях [167]:

А. После дескриптора, обозначающего химический эле
(присвоим ему символ N), даются дескрипторы N, ИЗОТОП
N, СПЛАВЫ (если N образует сплавы) и N, СОЕДИНЕН
(если N не инертный газ). Например:

ЛИТИЙ,
ЛИТИЙ, ИЗОТОПЫ
ЛИТИЙ, СПЛАВЫ
ЛИТИЙ, СОЕДИНЕНИЯ

---

[165] Atomkernenergie-Dokumentation beim Gmelin-Institut, Frankfurt/Main. AED-AB-1963-Suppl., Jan. 1963; AED-Information Service, Indexed Bibliography. Descriptor List. (Тезаурус составлен на английском языке).

[166] Еще две части — «Фирмы и институты» и «Географические названия» — непосредственно в «Тезаурус AED» пока не включены, хотя и используются при индексировании.

[167] Эти нарушения обусловлены особенностями английского языка («Тезаурус AED», как уже было указано выше, составлен на английском языке).

ЕМПЕРАТУРЫ &

в «Тезаурусе техни-
отмечены символами
тмеченный им деск-
сколькими другими
окие понятия и от-
отмечаются деск-
одного или не-
е узкие понятия.
ван в январе
ии при Институт
ит из следую-

ключает на-
ских соеди-
Предметы
ойств, эле-
ских про-
рзначения

перечис-
каждой
одном
букт-
троен-
ко в

нт
I:
Н

На английском языке этот пример им

LUTHIUM
LUTHIUM ISOTOPES
LUTHIUM ALLOYS
LUTHIUM COMPOUND

В названиях неорганических соединений
окси-, сульфо-, тио- и т. п. не учитываются
такие соединения даются в виде соответству
названий классов, к которым эти соединени
звания всех органических соединений даны
рядке под «заглавным» дескриптором ОРГАНИ
НЕНИЯ.

Б. В «Тезаурусе AED», кроме составных дес
CTOR CONTROL RODS (Управляющие стержни
CTOR SHIELDS (Защитные экраны реактора), RE
NOLOGY (Реакторная технология) и т. д., употре
дескрипторы, образованные путем добавления к слов
(Реакторы) соответствующих пояснительных сло
REACTORS, BREEDING RATIO (Реакторы, коэффи
изведения ядерного топлива), REACTORS, BUCKL
торы, лапласиан) и т. д. Дескрипторы второго типа
русе AED» даются после всех составных дескрипто
нающихся со слова REACTORS.

В «Тезаурусе AED» связи «заглавных» дескрипто
родовыми дескрипторами никак не обозначены. В словар
тьях этого тезауруса приводятся лишь видовые дескр
стоящие непосредственно ниже «заглавного» дискриптора.
они даются в алфавитном порядке под заголовком «смотри
причем для каждого видового дескриптора этот заголовок
ряется. Например:

ЭЛЕКТРОДЫ
— смотри также АНОДЫ
— смотри также КАТОДЫ

В некоторых словарных статьях «Тезауруса AED» приводятс
не все, а только некоторые из содержащихся в нем видовых деск
рипторов. Эти дескрипторы служат для целей ориентировки и да
ются в алфавитном порядке под одной общей рубрикой «смотри
также виды. . .» (далее следует название класса предметов или
явлений, обозначенных «заглавным» дескриптором). Например:

ЭЛЕМЕНТАРНЫЕ ЧАСТИЦЫ
— смотри также виды элементарных частиц
например: АЛЬФА-ЧАСТИЦЫ
МЕЗОНЫ

се AED» содержатся также дескрипторы НЕЙТ-РОНЫ, ПОЗИТРОНЫ, ПРОТОНЫ, ЭЛЕКТ-Однако эти дескрипторы, как видно из приведен-мера, опущены в словарной статье ЭЛЕМЕНТАР-Ы.

некоторых словарных статьях «Тезауруса AED» азание, как найти соответствующие видовые деск-ии видовые дескрипторы вообще не приводятся.

АЛИЗ
— смотри также ОТДЕЛЬНЫЕ МЕТОДЫ

отметить, что в «Тезаурусе AED» видовые деск-ют обратных отсылок к соответствующим родо-ам.

ками «смотри также» в словарных статьях «Тезау-водятся (в алфавитном порядке) и другие ассо-рипторы, которые в тезаурусе — в отличие от ви-ров — связаны между собой перекрестными ссыл-:

ПЕНИЕ
— смотри также ТЕПЛОПЕРЕДАЧА
. . . . . .
НВЕКЦИЯ
— смотри также ТЕПЛОПЕРЕДАЧА
. . . . . .
ПЛОПЕРЕДАЧА
— смотри также КИПЕНИЕ
— смотри также КОНВЕКЦИЯ
ЗМЕРЕНИЯ ТЕМПЕРАТУРЫ
— смотри также ПИРОМЕТРИЯ *[ассоциативный дескриптор]*
— смотри также виды устройств
напр.: ПИРОМЕТРЫ *[видовой дескриптор]*
РЕЗИСТИВНЫЕ ТЕРМОМЕТРЫ *[видовой дескриптор]*
. . . . . .
РОМЕТРИЯ
- смотри также ИЗМЕРЕНИЯ ТЕМПЕРАТУРЫ
. . . . . .
РОМЕТРЫ
. . . . . .
ИСТИВНЫЕ ТЕРМОМЕТРЫ

ятся примеры построения словарных статей в
168.

аты из второй части тезауруса.

1) БИОЛОГИЯ
— смотри также ЖИВОТНЫЕ
— смотри также ПРОДУКТЫ ПИТАНИЯ } *Другие ассоциативные дескрипторы*
— смотри также РАСТЕНИЯ
— смотри также отдельные отрасли
напр.: ГЕНЕТИКА
МИКРОБИОЛОГИЯ } *Видовые дескрипторы*
МОРСКАЯ БИОЛОГИЯ

2) ИОННЫЙ ОБМЕН
ионообменная хроматография *Синоним*
— смотри ХРОМАТОГРАФИЯ
ИОНООБМЕННИКИ
— смотри также ИОНООБМЕННИКИ. Часть 1
ИОНИЗАЦИЯ
ИОНИЗАЦИОННЫЕ ПОТЕНЦИАЛЫ
ИОНИЗАТОРЫ
ионизирующее излучение *Синоним*
— смотри виды ионизирующих излучений
напр.: АЛЬФА-ЧАСТИЦЫ } *Видовые дескрипторы*
ГАММА-ЛУЧИ
ИОНОСФЕРА

Из примера (2) видно, что некоторые словарные статьи разных частей «Тезауруса AED» связаны перекрестными ссылками.

6. «Тезаурус Евратома» (изд. 2)[169] используется в документальной ИПС Евратома, реализованной при помощи ЭЦВМ типа ИБМ-360/40. В предисловии к этому тезаурусу сказано, что он «служит справочным пособием (reference tool) для индексаторов, а также для документалистов, занимающихся формулированием поисковых предписаний». Кроме того, данный тезаурус записан на магнитных дисках и используется для автоматического контроля за правильностью написания дескрипторов в поисковых образах документов и в поисковых предписаниях, а также для замены в их текстах ключевых слов дескрипторами.

«Тезаурус Евратома» (изд. 2) состоит из двух частей (выпусков):

1) словаря ключевых слов и дескрипторов,

2) собрания графических схем (60), в которых выражены парадигматические отношения между дескрипторами.

В свою очередь, словарь ключевых слов и дескрипторов имеет два раздела: а) алфавитный список дескрипторов общего характера и б) общий алфавитный список ключевых слов и дескрипторов. Для сравнения следует указать, что в первом издании (1964 г.) этот тезаурус имел 5 следующих разделов: 1. Алфавитный список дескрипторов общего характера. 2. Алфавитный

[169] Euratom-Thesaurus. Indexing terms used within Euratom's nuclear documentation system. EUR 500. e, (2nd ed., pt. 1). Brussels, Directorate «Dissemination for Information», Center for Information and Documentation, 1966.

список дескрипторов, обозначающих неорганические с
3. Алфавитный список дескрипторов, обозначающих
4. Общий словарь ключевых слов. 5. Собрание графиче
предметных классов дескрипторов. Ниже приведены д
личественные характеристики двух изданий «Тезауру
тома».

| | Изд. 1 (1964 г.) | Из (196 |
|---|---|---|
| Общее число дескрипторов, | 4470 | 46 |
| в том числе: | | |
| Число дескрипторов общего характера | 1230 | 119 |
| Число дескрипторов, обозначающих неорганические соединения | 1836 | 1760 |
| Число дескрипторов, обозначающих изотопы | 1404 | 1635 |
| Число дескрипторов, обозначающих сплавы | — | 71 |
| Число ключевых слов | 2113 | 14518 |

Алфавитный список дескрипторов общего характера (1199) содержит дескрипторов, обозначающих неорганические соеди ния, изотопы и сплавы.

Общий алфавитный список ключевых слов и дескриптор содержит 3466 дескрипторов и 14518 ключевых слов. Каждом дескриптору в этом списке предшествует число, которое показывает, сколько раз данный дескриптор был использован для индексирования 360 тыс. документов, введенных в ИПС до сентября 1966 г. Такие числа дают хорошее представление о степени употребляемости каждого дескриптора. Они также позволяют оценивать объем выдачи, которую можно ожидать при той или иной формулировке поискового предписания. В качестве дескрипторов используются однозначные термины, которые обычно выражают простые понятия и имеют достаточно общий (неспецифичный) характер, гарантирующий их употребляемость в течение сравнительно большого периода времени.

Все ключевые слова, включенные во второе издание «Тезауруса Евратома», разделены на две категории: а) «приемлемые» (accepted) термины (11030) и б) «запрещенные» термины (3488). «Приемлемыми» ключевыми словами считаются:

— термины, являющиеся частичными синонимами дескрипторов;

— термины, которые по смыслу соответствуют понятиям, выражаемым сочетаниями дескрипторов;

— термины, выражающие слишком узкие понятия;

— имена собственные, используемые при обозначении теорий и методов;

— условные названия и обозначения проектов, реакторов, сплавов, химических соединений, минералов, биологических видов, некоторых вычислительных машин, а также применяемых в этих машинах кодов.

«Приемлемые» ключевые слова разрешается испол
индексировании наряду с дескрипторами (для облег
цедуры формулирования поисковых образов докуме
исковых предписаний). Однако в сформулированных
образах и предписаниях все ключевые слова заменяются
или индексатором) соответствующими дескрипторами.
ление «приемлемых» ключевых слов в функции дескрип
пускается лишь в исключительных случаях.

К категории «запрещенных» ключевых слов относя
мины, которые являются синонимами или аббревиату
скрипторов или «приемлемых» ключевых слов. В обще
витном списке перед каждым таким термином проставле
запрещающее использование этого термина при индексиро

В общем алфавитном списке дескрипторов и ключевы
«Тезауруса Евратома» (изд. 2) применяются указательны
меты (ссылки) двух типов:

1) «используй» (или «используй . . . +»);
2) «смотри» (или «смотри . . . или»).

Пометы первого типа требуют обязательного включения
исковый образ или предписание всех терминов, перечисле
против данного ключевого слова или дескриптора (в том ч
и замены «запрещенного» ключевого слова соответствующим
скриптором или сочетанием дескрипторов). Пометы второго т
указывают лишь на то, что в поисковом образе или предписан
может быть использован один или несколько дескрипторов,
торые перечислены против данного ключевого слова или дескри
тора.

В числе терминов, на которые делаются ссылки в «Теза
русе Евратома» (изд. 2), могут быть не только дескрипторы
но и «приемлемые» ключевые слова.

Ниже в качестве примера приведен фрагмент из «Тезауруса
Евратома» (изд. 2). В этот фрагмент включены все дескрипторы,
на которые сделаны ссылки в приводимых словарных статьях.
Например, «приемлемое» ключевое слово ЗАКОН БЕРНУЛЛИ
при индексировании должно быть заменено дескрипторами 18
ГИДРОСТАТИКА, 25947 ДАВЛЕНИЕ и 10490 ПОТОК ЖИД-
КОСТИ. Кроме того, дескриптор 18 ГИДРОСТАТИКА должен
быть дополнен дескрипторами 11089 ЖИДКОСТИ и 3433 МЕХА-
НИКА. Если после пометы «используй», стоящей против «запре-
щенного» ключевого слова, стоит символ «=», то это означает, что
данное ключевое слово считается несущественным.

17 АЗИЯ

2 АКТИНОЛИТ

ИСПОЛЬЗУЙ МИНЕРАЛЫ
\+ СИЛИКАТЫ ЖЕЛЕЗА
\+ СИЛИКАТЫ КАЛЬЦИЯ
\+ СИЛИКАТЫ МАГНИЯ

[illegible] ИСПОЛЬЗУЙ БАКТЕРИИ
[illegible] ИСПОЛЬЗУЙ АНТИБИОТИКИ
[illegible] ИСПОЛЬЗУЙ РАДОН 219

[illegible] СМОТРИ НЕРВНАЯ СИСТЕМА

[illegible] ИСПОЛЬЗУЙ ХРОНОТРОНЫ

[illegible] СМОТРИ РЕГЕНЕРАЦИЯ
ИЛИ РЕКРИСТАЛЛИЗАЦИЯ
ИЛИ РЕЛАКСАЦИЯ

[illegible] ИСПОЛЬЗУЙ МАГНИТНЫЕ ПОЛЯ
+ ТЕПЛОПЕРЕНОС
+ ЭЛЕКТРИЧЕСКИЕ ТОКИ

[illegible] ИСПОЛЬЗУЙ КОЛЕБАНИЯ
+ СХЕМЫ

2) [illegible] ИСПОЛЬЗУЙ ЖИДКОСТИ
+ МЕХАНИКА

[illegible] ДАВЛЕНИЕ
[illegible] ЖИДКОСТИ
ЗАКОН БЕРНУЛЛИ ИСПОЛЬЗУЙ ГИДРОСТАТИКА
+ ДАВЛЕНИЕ
+ ПОТОК ЖИДКОСТИ

3)
84323 ИЗМЕРЕНИЕ
9584 ИМПУЛЬСЫ
2629 ИНФРАКРАСНОЕ ИЗЛУЧЕНИЕ
— КНИГА ИСПОЛЬЗУЙ =
12465 КОЛЕБАНИЯ
23395 МАГНИТНЫЕ ПОЛЯ
4) 434 МАЗЕРЫ
3433 МЕХАНИКА
6977 МИНЕРАЛЫ
МУССОН ИСПОЛЬЗУЙ АЗИЯ
+ ВЕТЕР

11621 НАГРЕВ
2488 НЕРВНАЯ СИСТЕМА
НИКЕЛИНОВЫЙ СПЛАВ
НИКЕЛИНОВЫЙ СПЛАВ ИСПОЛЬЗУЙ СПЛАВЫ МЕДИ
+ СПЛАВЫ НИКЕЛЯ

10522 ПЕРЕХОДНЫЕ ПРОЦЕССЫ
19114 ПЛАЗМА
10490 ПОТОК ЖИДКОСТИ
6723 ПРОГРАММИРОВАНИЕ
32 РАДОН 219
— РАЗВИТИЕ — ИСПОЛЬЗУЙ =
1656 РЕГЕНЕРАЦИЯ
1097 РЕКРИСТАЛЛИЗАЦИЯ
1322 РЕЛАКСАЦИЯ — СМОТРИ ГЕНЕРАТОРЫ
ИЛИ КОЛЕБАНИЯ
ИЛИ МАЗЕРЫ
ИЛИ ПЕРЕХОДНЫЕ ПРОЦЕССЫ
ИЛИ ПЛАЗМА
ИЛИ УРАВНЕНИЯ
ИЛИ ЭЛЕКТРОЛИТЫ
350 СИЛИКАТЫ ЖЕЛЕЗА
346 СИЛИКАТЫ КАЛЬЦИЯ
443 СИЛИКАТЫ МАГНИЯ
3034 СПЛАВЫ МЕДИ
7017 СПЛАВЫ НИКЕЛЯ
СУЛЬФАТЫ КСЕНОНА
9633 СХЕМЫ
41624 ТЕМПЕРАТУРА
1684 ТЕПЛОВОЕ ИЗЛУЧЕНИЕ
—ТЕПЛОВОЙ ПОТОК — ИСПОЛЬЗУЙ ТЕПЛОПЕРЕНОС
—ТЕПЛОЕМКОСТЬ — ИСПОЛЬЗУЙ УДЕЛЬНАЯ ТЕПЛОЕМКОСТЬ
2509 ТЕПЛООБМЕННИКИ
8484 ТЕПЛОПЕРЕНОС
14902 ТЕРМОДИНАМИКА
—ТЕПЛОТА — СМОТРИ ИНФРАКРАСНОЕ ИЗЛУЧЕНИЕ
ИЛИ НАГРЕВ
ИЛИ ТЕМПЕРАТУРА
ИЛИ ТЕПЛОВОЕ ИЗЛУЧЕНИЕ
ИЛИ ТЕРМОДИНАМИКА
ИЛИ УДЕЛЬНАЯ ТЕПЛОЕМКОСТЬ
ИЛИ ЭНЕРГИЯ
2593 УДЕЛЬНАЯ ТЕПЛОЕМКОСТЬ
21543 УРАВНЕНИЯ

11800 ЭЛЕКТРИЧЕСКИЕ ТОКИ
1676 ЭЛЕКТРОЛИТЫ
41602 ЭНЕРГИЯ
18-0-КОД ИСПОЛЬЗ

*В «Тезаурусе Евратома» (изд. 2) применен вила построения дескрипторов, обозначающи единения, минералы, сплавы и изотопы:*

*А. Названия химических соединений выража восочетаний, образованных из названия соотв мического элемента (дескриптора) и другого де того из такого списка:*

| | |
|---|---|
| СОЕДИНЕНИЯ | НИТРАТ |
| СОЕДИНЕНИЯ КОМПЛЕКСНЫЕ | НИТРИД |
| БОРИДЫ | ОКИСЛЫ |
| БРОМИДЫ | СИЛИКАТ |
| ГИДРИДЫ | СУЛЬФАТЫ |
| ГИДРООКИСЛЫ | СУЛЬФИДЫ |
| ИОДИДЫ | ФОСФАТЫ |
| КАРБИДЫ | ФТОРИДЫ |
| КАРБОНАТЫ | ХЛОРИДЫ |

Например: БАРИЙ+ОКИСЛЫ=ОКИСЛЫ БАРИЯ
МАГНИЙ+СОЕДИНЕНИЯ=СОЕДИНЕНИЯ МАГН
СЕРА+ХЛОРИДЫ=ХЛОРИДЫ СЕРЫ

*При построении дескрипторов по этому правилу необ учитывать, что некоторые элементы не образуют всех видов ческих соединений, перечисленных в приведенном выше с Например, дескриптор ФТОР не сочетается с дескриптором РИДЫ, дескриптор СЕРА — с дескриптором КАРБИДЫ*

*Б. Состав комплексных химических соединений и минер может быть выражен путем сочетания нескольких дескрипто обозначающих компоненты соответствующего соединения или нерала. В частности, состав органических соединений реком дуется выражать в виде сочетаний дескрипторов, которые о значают следующие компоненты таких соединений: а) основ соединение; б) функциональные группы; в) замещающие групп Например:*

ДИФТОРЭТАНОЛ=ЭТАН *(основное соединение)*
+АЛЬДЕГИДЫ *(функциональная группа)*
+ФТОРИРОВАННЫЕ УГЛЕВОДОРОДЫ
*(замещающая группа)*

ЛЬЗУЙ ВРЕМЯ
+ ИЗМЕРЕНИЕ
+ ИМПУЛЬСЫ

УЙ ПРОГРАММИРО-
ВАНИЕ

ы следующие пра-
химические со-

ются в виде сло-
тствующего хи-
криптора, взя-

ы
ы

ы

ия

ходимо
хими-
иске.
ото-
т. д.
лов
ов,
и-
-

*В. Названия сплавов переводятся на ИП*
*четаний, образуемых из названия соответствую*
*талла) и дескриптора СПЛАВЫ. Кроме того,*
*может быть выражено в виде перечня дескрипто*
*значают существенные компоненты этого сплава.*
*в сплавах неметаллы выражаются дескрипторам*
*значают соединения, образуемые этими неметалл*
*дескрипторов, выражающие стехиометрические с*
*обходимо дополнять дескриптором ИНТЕРМЕТА*
*СОЕДИНЕНИЯ. Например:*

АНТИМОНИД ГАЛЛИЯ=СПЛАВЫ ГАЛЛИЯ
+СПЛАВЫ СУРЬМЫ
+ИНТЕРМЕТАЛЛИЧЕСКИЕ СО

*Г. Названия изотопов выражаются в виде слов*
*образуемых из названий соответствующих химическ*
*тов и их массовых чисел или дескриптора ИЗОТОПЫ.*

БОР 8, МЕНДЕЛЕВИЙ 256, УГЛЕРОД 14,
ИЗОТОПЫ ИОДА И Т. Д.

*Это правило построения не распространяется на изото*
*рода, которым в данном ИПЯ соответствуют лишь деск*
*ВОДОРОД, ДЕЙТЕРИЙ и ТРИТИЙ.*

*Парадигматические отношения между дескрипторами в*
*урусе Евратома» (изд. 2) выражены в графической форме (*
*лиц)* [170]. *Пример построения такой таблицы приведен на р*

*В заключение данного раздела необходимо отметить, что*
*заурусе Евратома» (изд. 2), как и во всех других тезаур.*
*все дескрипторы и ключевые слова, которые обозначают пере*
*лимые предметы и сочетаются с количественными числительны*
*даны во множественном числе. В «Тезаурусе Евратома» (изд*
*дескрипторы и ключевые слова набраны шрифтом одного кег*
*так как оригинал-макет тезауруса был напечатан выводным у*
*ройством ЭЦВМ типа ИБМ-360/40.*

*7. «Тезаурус Бюро кораблестроения и р*
*монта ВМФ США» (изд. 2) опубликован в марте 1965 г*
*(его первое издание вышло в декабре 1963 г.). Этот тезаурус*
*был разработан для автоматической ИПС, получившей название*
*Проект SHARP (SHip Analysis and Retrieval Project — Проект*
*анализа и поиска [документов] по кораблестроению) и реализо-*
*ванной при помощи ЭЦВМ типа ИБМ-7090 и ИБМ-1401* [171].

[170] *Эти таблицы еще не опубликованы.*

[171] *J. J. Nicolaus. The automated approach to technical information retrieval. Library applications. NAVSHIPS 250-210-2. Washington, Department of the Navy, Bureau of Ships, 1964; J. J. Nicolaus. Description of the Bureau of Ships Thesaurus of descriptive terms and code book. «Bull.*

«Тезаурус Бюро кораблестроения и ремонта ВМФ США» (далее мы будем называть его просто «Тезаурусом БКР») содержит свыше 4,6 тыс. дескрипторов. Вот как определяет назначение этого тезауруса Дж. Николос, заведующий технической библиотекой Командования военно-морских корабельных систем, в которой используется «Тезаурус БКР»: «Этот тезаурус представляет собой список ассоциированных слов (word-association list), которому придана такая иерархическая структура, чтобы индексаторы и предметизаторы могли при вводе документа в ИПС описывать его смысловое содержание с желаемой степенью широты (generality) или специфичности, а производящие поиск — описывать в точных терминах информацию, которую они желают получить на выходе. Это гибкий нормативный список, предназначенный для контроля за употребляемой терминологией и рассчитанный на использование терминов в сочетаниях друг с другом, что необходимо для координации понятий и чем он отличается от обычных списков предметных заголовков, которые не допускают такой степени гибкости. Тезаурус содержит только те термины, которые использованы для индексирования имеющегося в библиотеке собрания документов. По мере необходимости могут добавляться и новые термины, но они с помощью соответствующих перекрестных ссылок должны быть связаны с уже используемыми терминами» [172].

В «Тезаурусе БКР» применяются указательные пометы следующих типов:

а) includes (включает) — обозначает ключевые слова, считающиеся синонимами «заглавного» дескриптора. К категории синонимов отнесены устаревшие термины, разные варианты терминов и аббревиатуры дескрипторов. Например:

БАРОМЕТРИЧЕСКОЕ ДАВЛЕНИЕ
  ВКЛЮЧАЕТ:
    АТМОСФЕРНОЕ ДАВЛЕНИЕ
ЯДЕРНЫЕ РЕАКТОРЫ
  ВКЛЮЧАЕТ:
    АТОМНЫЕ КОТЛЫ
    АТОМНЫЕ РЕАКТОРЫ
    КОТЛЫ (АТОМНЫЕ)
    РЕАКТОРЫ (АТОМНЫЕ)
    РЕАКТОРЫ (ЯДЕРНЫЕ)

---

de L'AID», 1966, v. 5, № 4, стр. 59—65; Thesaurus of descriptive terms and code book. (NAVSHIPS 0900-002-0000). 2nd ed. Washington, Department of the Navy, Bureau of Ships, Technical Library, 1965.

[172] J. J. Nicolaus. The automated approach to technical information retrieval. Library applications. NAVSHIPS 250-210-2. Washington, Department of the Navy, Bureau of Ships, 1964, стр. 5—6.

б) use (используй) — обозначает дескрипторы, синонимом которых в данной ИПС считается «заглавное» ключевое слово. Например:

АТМОСФЕРНОЕ ДАВЛЕНИЕ
ИСПОЛЬЗУЙ:
БАРОМЕТРИЧЕСКОЕ ДАВЛЕНИЕ

в) broader terms (более широкие термины) — обозначает родовые дескрипторы. Например:

БИНОКЛИ
БОЛЕЕ ШИРОКИЕ ТЕРМИНЫ:
ОПТИЧЕСКИЕ ПРИБОРЫ

г) related terms (более узкие термины) — обозначает видовые дескрипторы. Например:

БИНОКЛИ
БОЛЕЕ УЗКИЕ ТЕРМИНЫ:
ИК-БИНОКЛИ

д) related terms (родственные термины) — обозначает дескрипторы, которые связаны с «заглавным» дескриптором лишь отношениями общего характера, «образующимися в результате частого совместного использования дескрипторов или их широких и/или ограниченных ассоциаций» [173]. Например:

РЕФЕРИРОВАНИЕ
РОДСТВЕННЫЕ ТЕРМИНЫ:
ДОКУМЕНТАЦИЯ
ПРЕДМЕТНОЕ ИНДЕКСИРОВАНИЕ
РЕФЕРАТЫ

Таким образом, мы видим, что по своей внутренней структуре «Тезаурус БКР» почти ничем не отличается от «Тезауруса технических терминов». Ниже в качестве примера приведен фрагмент из «Тезауруса БКР».

ЯДЕРНОЕ ИЗЛУЧЕНИЕ
(ИЗЛУЧЕНИЕ НЕЙТРОНОВ, ГАММА-КВАНТОВ
И ДРУГИХ ЧАСТИЦ ИЗ АТОМНОГО ЯДРА В РЕЗУЛЬТАТЕ
ДЕЛЕНИЯ ИЛИ СЛИЯНИЯ ЯДЕР)
ВКЛЮЧАЕТ:
ИЗЛУЧЕНИЕ ЯДЕРНОЕ
РОДСТВЕННЫЕ ТЕРМИНЫ:
ДЕЛЕНИЕ
ЯДЕРНЫЕ ЧАСТИЦЫ
. . . . . .

[173] J. J. Nicolaus. Description of the Bureau of Ships Thesaurus of descriptive terms and code book. «Bull. de l'AID», 1966, v. 5, № 4, стр. 63.

ЯДЕРНЫЕ СВОЙСТВА
(СВОЙСТВА, ХАРАКТЕРНЫЕ ДЛЯ ЯДЕР АТОМОВ, И ОБЩЕТЕОРЕТИЧЕСКИЕ ПОНЯТИЯ, СВЯЗАННЫЕ С ПОВЕДЕНИЕМ ЯДЕР)
БОЛЕЕ УЗКИЕ ТЕРМИНЫ:
СПИН ЯДРА
ЧЕТНОСТЬ
ЯДЕРНЫЕ МАГНИТНЫЕ МОМЕНТЫ
ЯДЕРНЫЕ МОМЕНТЫ
ЯДЕРНЫЕ ЭЛЕКТРИЧЕСКИЕ МОМЕНТЫ
ЯДЕРНЫЙ МАГНИТНЫЙ РЕЗОНАНС
ЯДЕРНЫЙ РЕЗОНАНС
РОДСТВЕННЫЕ ТЕРМИНЫ:
СТРОЕНИЕ ЯДРА
ЯДЕРНАЯ ФИЗИКА
ЯДЕРНЫЕ ЧАСТИЦЫ

«Тезаурус БКР» (изд. 2) состоит из двух томов: собственно тезауруса и кодовой книги. Кодовая книга имеет две части: (1) алфавитный список дескрипторов с присвоенными им буквенными кодами; (2) алфавитный список кодов с соответствующими им дескрипторами. В качестве кодов дескрипторов в Тезаурусе БКР используются семибуквенные сокращения этих дескрипторов, образуемые по следующим правилам: сначала идут первые три-четыре буквы данного дескриптора, за которыми следуют согласные и/или гласные, в зависимости от длины дескриптора. Предполагается, что построенные таким образом коды обладают мнемоничностью, которая облегчает контроль за правильностью написания дескрипторов в поисковых образах и предписаниях. Например:

ЛАМПЫ БЕГУЩЕЙ ВОЛНЫ=ЛАМПБВЛ
РАДИОЛОКАТОР =РАДЛКТР
МАГНИТЫ =МАГНИТЫ и т. д.

Из следующих изданий «Тезауруса БКР» кодовая книга будет исключена, так как тезаурус введен в ЭЦВМ. Это делает ненужным кодирование дескрипторов, вследствие чего становится ненужной и кодовая книга.

8. «Нормативный список предметных заголовков Американского нефтяного института» (изд. 3) [174] опубликован в январе 1966 г. (первое издание вышло в 1964 г.). Этот тезаурус (условимся называть его

[174] E. H. Brenner and J. G. Mulvihill. American Petroleum Institute information retrieval project: Subject authority list. «Bull. de l'AID», 1966, v. 5, №4, стр. 81—84; Subject authority list. American Petroleum Institute information retrieval project. 3 ed. New York, American Petroleum Institute, Central Abstracting Service, 1966.

«Тезаурусом АНИ») используется в ИПС, реализованной при помощи вычислительной машины. При составлении «Тезауруса АНИ» был применен фасетный анализ, в котором использовались следующие категории: *Материалы*, *Процессы*, *Явления*, *Оборудование*, *Живые организмы*, *Экономический фактор*, *Место*, *Свойство*, *Рабочее условие*, *Общие атрибуты*.

«Тезаурус АНИ» состоит из двух частей: а) алфавитного словаря дескрипторов и ключевых слов; б) систематического указателя дескрипторов (сгруппированы по перечисленным выше категориям), который выполняет функцию путеводителя по тезаурусу.

В «Тезаурусе АНИ» применяются следующие виды указательных помет:

а) use, use . . . or, use . . . plus (используй, используй . . . или, используй . . . плюс) — обозначает один или несколько дескрипторов, синонимом которых в данной системе считается «заглавное» ключевое слово;

б) broader terms (autoposted) более широкие термины (добавляются автоматически) — обозначает родовые дескрипторы, которыми ЭЦВМ автоматически дополняет поисковый образ или предписание, если оно содержит данный «заглавный» дескриптор;

в) narrower terms (более узкие термины) — обозначает видовые дескрипторы;

г) see also (смотри также) — обозначает другие дескрипторы, ассоциативно связанные с «заглавным» дескриптором. Если после дескриптора, расположенного под заголовком «смотри также», проставлена звездочка, то такой дескриптор автоматически вводится в поисковый образ или предписание, содержащее данный «заглавный» дескриптор;

д) used for (используется вместо) — обозначает в словарной статье ключевые слова, которые считаются синонимами «заглавного» дескриптора.

Ниже в качестве примера приведены фрагменты из обеих частей «Тезауруса АНИ».

1) РАДИАЦИОННАЯ СТОЙКОСТЬ
ИСПОЛЬЗУЙ
РАДИАЦИЯ ПЛЮС УСТОЙЧИВОСТЬ
ИЛИ
РАДИАЦИЯ ПЛЮС ПРОНИЦАЕМОСТЬ
РАДИАЦИОННАЯ ФИЗИКА
ИСПОЛЬЗУЙ
ЯДЕРНАЯ ФИЗИКА ПЛЮС РАДИАЦИЯ
РАДИКАЛ СВОБОДНЫЙ
ИСПОЛЬЗУЙ
СВОБОДНЫЙ РАДИКАЛ

РАДИОАКТИВНОСТЬ
БОЛЕЕ ШИРОКИЕ ТЕРМИНЫ (ДОБАВЛЯЮТСЯ АВТОМАТИЧЕСКИ)
РАДИАЦИЯ
БОЛЕЕ УЗКИЕ ТЕРМИНЫ
ПЕРИОД ПОЛУРАСПАДА
СМОТРИ ТАКЖЕ
ДЕЙТЕРИЙ
КОБАЛЬТ 60*
ТРИТИЙ*
УГЛЕРОД 14*
ИСПОЛЬЗУЕТСЯ ВМЕСТО
ЯДЕРНОЕ ИЗЛУЧЕНИЕ
РАДИОЛИЗ
ПОЯСНЕНИЯ
ДЛЯ БОЛЕЕ ПОЛНОГО ПОИСКА
ИЩИ РЕАГЕНТЫ И ПРОДУКТЫ
СМОТРИ ТАКЖЕ
БОМБАРДИРОВКА
ИОНИЗАЦИЯ
ФОТОЛИЗ
2) ОБОРУДОВАНИЕ (ПРОДОЛЖЕНИЕ)
ИСКУССТВЕННЫЙ СПУТНИК
КОРАБЛЬ
БАРЖА
ГИБКАЯ БАРЖА
БУКСИР
ТАНКЕР
ТАНКЕР С КОРМОВЫМ МОСТИКОМ
ТАНКЕР-РЕФРИЖЕРАТОР
КОСМИЧЕСКИЙ КОРАБЛЬ
РАКЕТА
НАВИГАЦИОННОЕ ОБОРУДОВАНИЕ
САМОЛЕТ
УПРАВЛЯЕМЫЙ СНАРЯД

9. «Тезаурус терминов по целлюлозе и бумаге» опубликован Канадским научно-исследовательским институтом целлюлозы и бумаги в марте 1965 г. В нем содержится около 2 тыс. дескрипторов. Тезаурус по структуре ничем не отличается от «Тезауруса технических терминов» [175].

10. «Тезаурус к «Engineering Index» (США) [176] начал составляться еще в 1964 г., но до сих пор не опубликован.

[175] P. M. Nobbs. Pulp and Paper Research Institute of Canada. «Bull. de l'AID», 1966, v. 5, № 4, стр. 79—80; Thesaurus of pulp and paper terms. Montreal, Pulp and Paper Research Institute of Canada, 1965.

[176] F. R. Whaley and C. M. Flanagan. The Engineering Index Thesaurus. «Bull. de l'AID», 1966, v. 5, № 4, стр. 45—52.

Этот тезаурус предназначен для использования в автоматической информационной системе, которая создается фирмой Engineering Index, Inc. и названа системой CADRE (Current Awareness and Document retrieval for Engineers — Текущее оповещение и поиск документов для инженеров). В настоящее время в тезаурусе насчитывается свыше 11 тыс. дескрипторов. По внутренней структуре этот тезаурус практически ничем не отличается от «Тезауруса технических терминов».

«Тезаурус к «Engineering Index» используется также для машинного составления алфавитно-предметных указателей, причем дескрипторы выступают в роли предметных заголовков. Если в словарной статье дескриптора, использованного для индексирования какого-либо документа, имеются дескрипторы, которые отмечены аббревиатурами UF, BT и RT со звездочками, то при составлении указателя вычислительная машина, наряду с рубрикой [«заглавный» дескриптор — заглавие документа + шифр документа], автоматически порождает также рубрики [дескриптор BT * или RT * — смотри также «заглавный» дескриптор] или рубрики [дескриптор — смотри «заглавный» дескриптор]. Указательные пометы use, UF, BT, NT и RT здесь имеют такое же значение, что и в «Тезаурусе технических терминов».

11. «Тезаурус черной металлургии» опубликован в мае 1966 г. женевским отделением итало-швейцарской фирмы «Centro sperimentale metallurgico S. p. A.» [177]. Этот тезаурус представляет собой лишь итальянский или англо-итальянский словарь дескрипторов и ключевых слов по черной металлургии. Отбор ключевых слов для этого тезауруса производился на основании частоты их употребления в годовых указателях к журналу «Journal of the Iron and Steel Institute» за 1962—1965 гг.

Тезаурус состоит из двух разделов: а) алфавитного списка 1340 дескрипторов (итальянского или англо-итальянского), в котором дескрипторы, выражающие общие понятия, отмечены символом °; б) словаря (английского или итальянского) ключевых слов, каждому из которых поставлен в соответствие один или несколько дескрипторов. Например:

| | | |
|---|---|---|
| БЕССЕМЕРОВА СТАЛЬ | *смотри* | КОНВЕРТЕРНАЯ СТАЛЬ |
| БЕТАТРОН | *смотри* | РАДИОГРАФИЯ *и* БЕТА-ЛУЧИ |
| ПИРОМЕТАЛЛУРГИЯ | *смотри* | ПЛАВЛЕНИЕ *или* ОБЖИГ *или* |

[177] V. Canepa e M. Trigari. Thesaurus. Copia di lavoro in edizione anglo-italiana. Genova, Centro sperimentale metallurgico S. p. A., 1966; Thesaurus per la siderurgia. Roma-Genova, Centro sperimentale metallurgico — S. p. A., 1966—67; A. Palazzi, V. Canepa, L. Pompilio, M. G. Trigari. Contributo sperimentale al reperimento automatico della documentazione siderurgica. «La metallurgia italiana», 1967, № 2, стр. 132—148.

РАДИОХИМИЯ *смотри* ДИСТИЛЛЯЦИЯ
ЯДЕРНАЯ ФИЗИКА *и/или*
РАДИОАКТИВНОСТЬ

12. «Тезаурус по диффузии» составлен П. Германном и Г. Лёшнером в Научно-исследовательском институте цветных металлов (г. Фрейберг, ГДР) [178] и опубликован в 1967 г. В тезаурусе насчитывается 185 дескрипторов и 291 ключевое слово. Тезаурус состоит из двух разделов — систематического и алфавитного.

В систематическом разделе тезауруса словарные статьи дескрипторов и ключевых слов сгруппированы в 9 классов, причем класс 1 «Вещества» делится еще на 3 подкласса.

Ниже в качестве примера приведен фрагмент из этого раздела тезауруса.

1. 1. Вещества вообще
ODE Газ
*используй лишь в:*
1. Диффузия в газах
2. Диффузия газов вообще
3. Диффузия отдельных газов *кроме:*
Азот
Водород
Инертные газы
Кислород
Фтор
Хлор
ND Минерал
D Прочие неорганические вещества
D Органические вещества
D Радиоизотоп
AD Активационный анализ
или AD Устойчивый изотоп
ND Шлак
D Полупродукт

Точно такую же структуру имеют словарные статьи и в алфавитной части тезауруса.

В тезаурусе по диффузии применены следующие указательные пометы:

а) D (Deskriptor) — дескриптор;

б) ND (Nicht-Deskriptor) — недескриптор, ключевое слово;

в) ODE (Oberdeskriptor mit eingeschränktem Begriffsumfang) — вышестоящий дескриптор, выражающий понятие ограниченного объема, которое образовано путем объединения подчиненных ему понятий (дескрипторов). Такие дескрипторы используются лишь в тех случаях, когда при индексировании нельзя ограничиться применением подчиненных им дескрипторов;

[178] P. Herrmann und G. Löschner. Diffusions-Thesaurus. Freiberg, Forschungsinstitut für NE-Metalle, 1967.

г) AD (Assoziierter Deskriptor) — ассоциативный дескриптор. Это «нормативный термин, выражающий понятие, которое пересекается с понятиями, выражаемыми другими терминами, что обусловлено наличием в содержании этих понятий хотя бы одного общего свойства или признака. Такое пересечение может быть причиной возникновения следующих видов ассоциаций: по смежности во времени и пространстве; по сходству; по контрасту; причинно-следственных; целое — часть и т.д. [179];

д) AND (Assoziierter Nicht-Deskriptor) — ассоциативный недескриптор. Эта помета используется для обозначения ключевых слов, которые в указанном выше отношении ассоциируются с другими терминами;

е) AODE (Assoziierter Oberdeskriptor mit eingeschränktem Begriffsumfang) — ассоциативный вышестоящий дескриптор, выражающий понятие ограниченного объема.

В заключение данного раздела приводится список некоторых других опубликованных дескрипторов по разным отраслям науки и техники.


1. Словарь дескрипторов по химии и химической промышленности. В. Б. Маргаритов и В. К. Шелученко (отв. ред.). М., Научно-исслед. ин-т техн.-экон. исследований М-ва хим. пром-сти СССР, 1966 [переплет 1967).
2. H. Buntrock. Thesaurus of documentation (Provisional descriptor list). Ispra, EURATOM/CETIS, 1964.
3. ECAC Thesaurus of keywords (AD-439204). Electromagnetic Compatibility Analysis Center, 1963.
4. W. Hammond and S. Rosenborg. Indexing terms of announcement publications for government scientific and technical research reports. A composite vocabulary. V. 1. Alphabetic listing (PB-181-712). V. 2. Structural listing (Technical report No. IR 15) (PB-181-713). Silver Spring, Md., Datatrol Corp., 1964.
5. C. K. Schultz. Information science thesaurus (working paper). Philadelphia, Drexel Institute of Technology, 1964.
6. TDCK-Circular thesaurus system. 4th ed., Netherlands Armed Forces Technical Information Center, 1966.
7. Tezaur pentru indexarea coordonata in domeniul documentării şi bibliotecomiei. Bucureşti, Institutul de documentare tehnică, 1965.
8. Thesaurus for central information reference and control. Wrigth-Patterson, Air Force Base, Foreign Technology Division, 1965.
9. A Thesaurus of civil defense descriptors (AD-611438). Durham, N. C., Research Triangle Institute, 1965.
10. Thesaurus of FAA descriptors (AD-604408). Washington, D. C., Federal Aviation Agency, 1964.
11. Thesaurus of terms on copper technology. N. Y., Copper Development Association, 1965.
12. Technical information thesaurus. Poughkeepsie, New York, IBM Data Systems Division, Development Laboratory, 1964.
13. M. H. Weik and V. J. Confer. Thesaurus of the names of electronic digital computers and data processors. Aderdeen Proving Ground, Md., Ballistic Research Laboratories, 1962.


[179] Там же, стр. 3.

## ОБОБЩЕННАЯ СТРУКТУРНАЯ СХЕМА ТЕЗАУРУСА

Проведенный в предыдущем разделе анализ важнейших зарубежных тезаурусов позволяет построить обобщенную структурную схему тезауруса, которая может быть описана следующим образом. По определению информационно-поисковый тезаурус — это нормативный словарь-справочник, предназначенный для индексаторов и потребителей информации. Тезаурус состоит из трех частей:

а) словарная часть, представляющая собой общий алфавитный список дескрипторов (вместе с их словарными статьями) и ключевых слов, которые в данной ИПС считаются синонимами этих дескрипторов;

б) «семантическая карта» словарного состава ИПЯ, представляющая собой систему смысловых классов, в которые сгруппированы все дескрипторы этого ИПЯ. В этом разделе тезауруса должны быть возможно более наглядно выражены важнейшие парадигматические отношения между дескрипторами или, по крайней мере, отношения «вид—род» («род—вид»);

в) руководство по переводу ключевых слов и словосочетаний естественного языка на дескрипторный ИПЯ.

А. В словарной части тезауруса дескрипторы и синонимичные им ключевые слова, как было указано выше, приводятся в виде общего алфавитного списка, причем дескрипторы в этом списке соответствующим образом отмечены (например, даны прописными буквами). В этом списке дескрипторы и их синонимы связаны перекрестными ссылками. Полисемия и омонимия дескрипторов и ключевых слов в ИПЯ устранены лексикографически, т. е. при помощи системы помет и специфических символов.

В данном разделе тезауруса каждый дескриптор образует словарную статью, которая обычно строится по следующей схеме:

$$d_i\langle M_{i1},\ M_{i2},\ M_{i3},\ M_{i4}\rangle,$$

где $d_i$ — «заглавный» дескриптор; $M_{i1}$ — упорядоченное по алфавиту множество ключевых слов, которые в пределах данной ИПС считаются синонимами $d_i$; $M_{i2}$ — упорядоченное по алфавиту множество дескрипторов, каждый из которых связан с $d_i$ отношением «род—вид»; $M_{i3}$ — упорядоченное по алфавиту множество дескрипторов, каждый из которых связан с $d_i$ отношением «вид—род»; $M_{i4}$ — упорядоченное по алфавиту множество дескрипторов, каждый из которых связан с $d_i$ одним из следующих парадигматических отношений: «целое—часть», «часть—целое», «причина—следствие» и функциональное сходство.

Любое из перечисленных множеств может быть одноэлементным и даже пустым. Множества $M_{i1}$, $M_{i2}$, $M_{i3}$, $M_{i4}$ в словарном разделе тезауруса обозначаются соответствующими пометами. Необходимо указать, что множество $M_{i1}$ в совокупности с $d_i$

образует класс условной эквивалентности, который и является дескриптором в строгом понимании этого термина. Однако здесь множество $M_{i1}$ выполняет функцию номинального определения, которое уточняет смысл $d_i$, выбранного для обозначения этого класса условной эквивалентности.

Словарную часть тезауруса целесообразно дополнить картотекой так называемых «неопределенных» ключевых слов. При переиздании тезауруса одни из этих ключевых слов могут быть введены в словарный состав ИПЯ, а другие — стать синонимами дескрипторов. «Неопределенные» ключевые слова иногда разрешается включать в поисковые образы документов.

Б. «Семантическая карта» словарного состава дескрипторного ИПЯ может быть выполнена в любом из двух следующих вариантов:

а) в виде сочетания множества тематических классов (полей) с множеством тематических групп дескрипторов. Тематические классы (поля) состоят из упорядоченных по алфавиту имен тематических групп, а тематические группы — из упорядоченных по алфавиту дескрипторов. В качестве имен тематических классов (полей) могут быть использованы слова и словосочетания естественного языка, не являющиеся дескрипторами, а в качестве имен тематических групп — один или несколько дескрипторов, входящих в данную группу. Эти два множества могут быть дополнены системой классов, в которые объединены дескрипторы, связанные между собой родо-видовыми отношениями.

б) в виде графических схем, в которых важнейшие парадигматические отношения между дескрипторами выражены при помощи соответствующих стрелок, как это сделано, например, в «Тезаурусе Евратома».

В. Правила перевода ключевых слов и словосочетаний естественного языка на дескрипторной ИПЯ регламентируют процедуру подстановки дескрипторов вместо ключевых слов и словосочетаний на заключительном этапе индексирования, а также описывают способы выражения в терминах данного ИПЯ таких ключевых слов и словосочетаний, которые вообще отсутствуют в словарном составе последнего или имеют в нем другую морфологическую форму. Здесь речь идет главным образом о переводе на дескрипторный ИПЯ названий многочисленных учреждений и организаций, сложных химических соединений, сплавов, минералов, биологических видов, исследовательских программ и т. п.

В правила перевода обычно включаются и правила лексикографического редактирования поисковых образов и предписаний перед их вводом в ИПС. Лексикографическое редактирование поисковых образов и предписаний заключается в их дополнении дескрипторами, которые связаны с основными дескрипторами родовидовыми и другими парадигматическими отношениями.

где $D_{i1}, D_{i2}, D_{i3}, \ldots, D_{ij}$ — предметные классы дескрипторов, образующих поисковый образ документа $I$; $a, b, c, \ldots, n$ — характеристики, которые соответственно показывают, что все дескрипторы, включенные в класс $D_i$, обозначают вещества, или свойства, или процессы и т. п. Из сказанного выше очевидно, что предметная группировка дескрипторов в поисковом образе есть не что иное, как обращенный вариант метода интерфиксов.

Основной недостаток указателей связи как весьма эффективного средства предотвращения «поискового шума» состоит в том, что при их использовании возможно снижение полноты выдачи. М. Таубе показывает это на следующем примере [185].

Допустим, что основным смысловым содержанием документа $N$ является покрытие медных труб свинцом. Тогда поисковый образ этого документа будет иметь следующий вид:

СВИНЕЦ, ПОКРЫТИЕ, МЕДЬ, ТРУБЫ.

При такой формулировке поискового образа ИПС будет выдавать документ $N$ в ответ на информационный запрос: «В каких документах говорится о свинцовых трубах?» Однако в документе $N$ ничего не говорится о свинцовых трубах, т. е. этот документ является нерелевантным по отношению к информационному запросу.

Для предотвращения выдачи документа $N$ на такой информационный запрос в поисковый образ этого документа могут быть введены соответствующие указатели связи. Тогда поисковый образ документа $N$ получит следующий вид:

СВИНЕЦ $R_1$, ПОКРЫТИЕ $R_1$, МЕДЬ $R_2$, ТРУБЫ $R_2$.

Если поставить условие, что документ должен выдаваться лишь в тех случаях, когда к дескрипторам поискового образа, совпадающим с дескрипторами поискового предписания, приписаны одинаковые указатели связи, то при таком построении поискового образа документ $N$ не будет выдаваться на информационный запрос: «В каких документах говорится о свинцовых трубах?». Однако тогда ИПС не будет выдавать документа $N$ и на информационный запрос: «В каких документах говорится о покрытии труб?».

Выдачу документа $N$ на такой информационный запрос, казалось бы, можно обеспечить путем введения в поисковый образ этого документа дополнительных указателей связи:

---

[185] M. Taube. Notes on the use of roles and links in coordinate indexing. «American Documentation», 1961, v. 12, № 2, стр. 98—100.

СВИНЕЦ $R_1R_3$, ПОКРЫТИЕ $R_1R_3$, МЕДЬ $R_2R_3$, ТРУБЫ $R_2R_3$.

Однако в этом случае не выполняется требование полного совпадения указателей связи, приписанных к дескрипторам ПОКРЫТИЕ и ТРУБЫ, и поэтому ИПС не выдает документа $N$ на приведенный выше информационный запрос.

Предположим, что это требование изменено и для выдачи документа достаточно совпадения хотя бы одного указателя связи из числа таких указателей, приписанных к соответствующим дескрипторам. Но тогда указатели связи вообще перестают выполнять свое назначение. Действительно, к каждому дескриптору в последнем варианте поискового образца документа $N$ приписан хотя бы один указатель связи, который одновременно приписан и к любому другому дескриптору этого поискового образа.

Наконец, возможна и такая формулировка поискового образа документа $N$:

СВИНЕЦ $R_1$, ПОКРЫТИЕ $R_1R_3$, МЕДЬ $R_2$, ТРУБЫ $R_2R_3$.

Если в качестве условия выдачи документа принять совпадение хотя бы одного указателя связи из числа таких указателей, приписанных к соответствующим дескрипторам поискового образа документа, то документ $N$ будет выдаваться на информационные запросы о покрытиях труб, о свинцовых покрытиях, о медных трубах и не будет выдаваться на информационный запрос о свинцовых трубах. Однако ИПС не будет выдавать этого документа и на информационный запрос о покрытиях меди, хотя в документе $N$ говорится об этом.

Теоретическое исследование данного вопроса показывает, что никакое усложнение системы указателей связи не может предотвратить неизбежных потерь релевантных документов при информационном поиске, если требуется обеспечить нулевой «поисковый шум». При этом М. Таубе подчеркивает, что указатели связи фактически не обозначают логических связей между дескрипторами поискового образа того или иного документа, а делят текст этого документа на отдельные части и относят соответствующие дескрипторы к этим частям. Поэтому ошибки возникают всякий раз, когда делается попытка использовать указатели такого типа для обозначения логических связей между дескрипторами, относящимися к одной и той же части текста.

Практический опыт свидетельствует, что наилучшие результаты достигаются в тех случаях, когда указатели связи применяются в сочетании с указателями роли. Какова же эффективность применения таких указателей? По ориентировочным данным, указатели роли и связи уменьшают «поисковый шум» на 10—15%. Однако их введение в дескрипторный ИПЯ увеличивает стоимость индексирования документов, причем пока нет никаких

данных о том, в какой степени этот рост расходов на индексирование компенсируется снижением «поискового шума».

Из других методов, применяемых в дескрипторных ИПЯ для предотвращения «поискового шума», упоминания заслуживают только следующие два. Это — образование словосочетаний из дескрипторов и введение таких словосочетаний в словарный состав ИПЯ в качестве самостоятельных лексических единиц. Этот метод отличается высокой эффективностью, но его практическое применение едва ли целесообразно, так как ведет к чрезмерному расширению словарного состава дескрипторного ИПЯ. Другой метод — построение *кортежей*, т. е. упорядоченных цепочек дескрипторов. При использовании этого метода поисковый образ каждого документа строится по следующей схеме:

$$P_i = d^1_{i3},\ d^2_{i1},\ d^3_{i2},\ \ldots,\ d^n_{ia},$$

где $P_i$ — поисковый образ документа $I$; $d_{i1}, d_{i2}, d_{i3}, \ldots, d_{ia}$ — дескрипторы, образующие $P_i$; 1, 2, 3, . . ., $n$ — порядковые номера соответствующих предметных классов дескрипторов в кортеже.

В данном случае синтагматические отношения между дескрипторами в поисковом образе выражаются через их последовательность в кортеже. Например, первый дескриптор в кортеже может обозначать вид действия, второй — место действия, третий — время действия и т. д. Если в поисковом образе документа не содержится дескриптора, принадлежащего к классу, за которым в кортеже закреплен порядковый номер $k$, то вместо него в кортеж вписывают нулевой дескриптор. Очевидно, что в кортеже каждый элемент может иметь только один предшествующий и один последующий дескриптор. Поэтому такой метод не позволяет выражать сложных синтагматических отношений между дескрипторами поискового образа.

В заключение данного раздела необходимо подчеркнуть, что введение в дескрипторный ИПЯ грамматических средств значительно усложняет процедуры индексирования и информационного поиска, делает их более дорогостоящими. Поэтому прежде чем ввести какое-либо усложнение в дескрипторный ИПЯ, необходимо тщательно изучить, будет ли в достаточной мере компенсировано увеличение стоимости индексирования и информационного поиска ростом его точности и полноты. Можно утверждать, что при сравнительно небольших поисковых массивах (примерно до 30 тыс. документов) повышение точности информационного поиска, достигаемое благодаря введению в дескрипторный ИПЯ соответствующих грамматических средств, обычно не компенсирует увеличения затрат времени и средств на индексирование и информационный поиск. Для небольших поисковых массивов оказывается выгодней применение дескрипторных ИПЯ без грамматики. Разумеется, многое здесь зависит и от специфичных требований, предъявляемых к той или иной ИПС.

## СПЕЦИАЛЬНЫЕ ИНФОРМАЦИОННО-ПОИСКОВЫЕ ЯЗЫКИ

Как уже было сказано, кроме дескрипторных языков, для поиска документальной информации широко применяются также библиотечно-библиографические классификации и алфавитно-предметные системы. Эти типы информационно-поисковых языков уже описаны в главах 3 и 5, и поэтому здесь они не рассматриваются. Однако в информационной практике используются или создаются более сложные формализованные ИПЯ, из которых наиболее известными являются Семантический код и язык Syntol.

Ниже дается краткое описание Семантического кода, разработанного под руководством известных американских специалистов в области научной информации Дж. Перри и А. Кента.

Алфавит Семантического кода состоит из следующих символов: прописных букв от *A* до *Z*, чисел от 000 до 999 и пунктуационных знаков *, -&¶. /.

Буквы, числа, косая черта и точка используются для построения слов, а остальные символы — для соединения слов в подфразы, фразы, предложения и параграфы.

Слова в Семантическом коде могут быть двух типов: указатели роли и дескрипторы. Указатели роли — это трехбуквенные слова вида *KOV.*, *KAP.*, *KAL.* и т. д., за каждым из которых следует точка. Все указатели роли начинаются с буквы *K*. Дескрипторы могут быть простыми и составными. Простой дескриптор всегда состоит из четырех букв: первая, третья и четвертая буквы обозначают *семантический фактор* (широкий класс предметов или понятий), а вторая буква — *инфикс*. Инфикс обозначается буквами *A*, *E*, *I*, *O*, *U*, *Q*, *W*, *X*, *Y*, *Z* и используется для построения дескриптора, обозначающего некоторый предмет или понятие, которое находится в определенном логическом отношении к другому предмету или понятию. Это можно показать на следующем примере. В Семантическом коде слову «рыба» соответствует дескриптор *FASH*. Семантический фактор *F-SH* обозначает класс «рыбы», а инфикс *A* — принадлежность к этому классу. Инфикс *Z* в Семантическом коде обозначает «быть похожим на что-либо». Если в дескрипторе *FASH* мы заменим *A* инфиксом *Z*, то получим новый дескриптор *FZSH.*, означающий «что-то похожее на рыбу», т. е. морское животное.

Из простых дескрипторов может быто построен составной дескриптор, в котором простые дескрипторы располагаются в углубленном алфавитном порядке (инфиксы при такой расстановке не учитываются). Методику построения составных дескрипторов путем изменения инфиксов в простых дескрипторах и соединения этих дескрипторов можно показать на следующем примере. Рассмотрим дескриптор *MACH.MUSR.RWHT*. Значения

дескрипторов *MACH.*, *MUSR.* и *RWHT.*, а также инфиксов *A*, *U* и *W* следующие:

| *Дескриптор* | *Обозначаемый предмет или понятие* | *Инфикс* | *Обозначаемое отношение* |
|---|---|---|---|
| *MACH.* | Машина или устройство | *A* | Индентичность |
| *MUSR.* | Мера, измерение | *U* | «Производит» |
| *RWHT.* | Тепло | *W* | «Испытывает действие на себе» |

На основании приведенных данных можно видеть, что дескриптор *MACH.MUSR.RWHT.* описывает машину или устройство (*MACH.*), которое *производит* (*U*) *измерение* (*MUSR.*) и *испытывает на себе действие* (*W*) *тепла* (*RWHT*). Таким образом, дескриптор *MACH.MUSR.RWHT.* означает «термометр».

Очевидно, что путем такого соединения простых дескрипторов можно придать составному дескриптору бóльшую специфичность, т. е. сузить класс предметов или понятий, обозначаемых этим дескриптором. Однако подобный метод в Семантическом коде не применяется. Вместо этого применяются цифровые суффиксы, которые можно приписывать к некоторым дескрипторам. Например, дескриптор *FZSH*.009. означает «тюлень», *FZSH*.014. — «кит», *FZSH*.017. — «дельфин» и т. д.

Поисковые образы документов формулируются в виде так называемых *рефератов телеграфного стиля*, которые затем переводятся на Семантический код. Реферат телеграфного стиля делится на один или несколько *параграфов*, параграф — на одно или несколько *предложений*, предложение — на одну или несколько *фраз*, фраза — на одну или несколько *подфраз*, а подфраза — на одну или несколько *подподфраз*. Каждая подподфраза состоит из одного или нескольких указателей роли, за которыми следует дескриптор.

Для выделения в реферате различных синтаксических структур используются соответствующие знаки пунктуации. Начало параграфа (кроме первого) и его конец обозначается символом ¶, начало и конец предложения (кроме первого) — символом &, начало и конец каждой фразы (кроме первой) — символом -, начало и конец каждой подфразы — символом [,].

Правила обозначения подфраз более сложны. Подфразе предшествует символ *, если выполнено одно из двух условий:

1) подфраза состоит более чем из одной подподфразы;

2) подфраза содержит такой же указатель роли, как предшествующая ей подфраза, и эта предшествующая подфраза содержит более одной подподфразы.

Если знаки пунктуации примыкают друг к другу, то они должны ставиться в следующем порядке:

, * - & ¶

Составляя реферат телеграфного стиля, референт стремится отразить в нем основные предметы, о которых говорится в документе, а также отношения между этими предметами. Ниже в качестве примера приводится запись предложения «Скорость старения и порчи бронзовых насосов от тепла» в Семантическом коде [186].

*KOV. FWLD. MACH. MQTN. PQRS.*001., *KUJ.LALL.* □ *CQU.* □ *SQN.*, *KWV.KAP.DADD.*001., *KAL.RAHT.*001., *KAL.GAPR.*041., *KAL.RAPR.TYMM.*001., — &

---

[186] J. W. Perry, A. Kent. Tools for machine literature searching. Semantic code dictionary. Equipment. Procedures. New York, Interscience Publishers, 1958, стр. 335; L. E. Allen, R. B. S. Brooks, P. A. James. Automatic retrieval of legal literature. Why and how. New Hawen, Meyer, 1962, стр. 36—40; J. L. Melton. The Semantic code today. «American Documentation», 1962, v. 13, № 2, стр. 176—181.

Глава 7

# МЕТОДЫ И СРЕДСТВА РЕАЛИЗАЦИИ ИНФОРМАЦИОННО-ПОИСКОВЫХ СИСТЕМ

## ОБЩИЕ ПРИНЦИПЫ

По определению, которое было дано в начале четвертой главы, в понятие абстрактной информационно-поисковой системы не были включены в качестве ее составной части специальные методы и средства реализации таких систем. И это вполне понятно, так как, с точки зрения ИПЯ, правила перевода с естественного языка на этот ИПЯ, а также критерий смыслового соответствия не зависят от того, где и как записываются поисковые образы документов, как осуществляется их считывание, какие методы применяются для проверки критерия смыслового соответствия, где и в каком виде хранится поисковый массив документов и т. п. Однако без использования таких методов и средств невозможно создание ни одной конкретной, реально действующей ИПС. Кроме того, от правильного выбора соответствующих методов и средств главным образом и зависит экономическая эффективность каждой конкретной ИПС, т. е. ее быстродействие и стоимость выдачи.

Для осуществления документального информационного поиска необходимы специальные методы и средства, обеспечивающие выполнение следующих операций:

а) запись поискового образа документа вместе с этим документом (или его копией) или вместе со сведениями, достаточными для его быстрой идентификации и отыскания;

б) запись поискового предписания;

в) считывание поисковых образов документов;

г) проверка критерия смыслового соответствия и принятие решения о выдаче или невыдаче документа;

д) выдача документов, отвечающих на информационный запрос, или соответствующих сведений о таких документах.

Совокупность различных средств, которые обеспечивают выполнение перечисленных операций, образует ИПС в широком понимании этого слова, т. е. систему, которая используется для практического поиска документальной информации. На

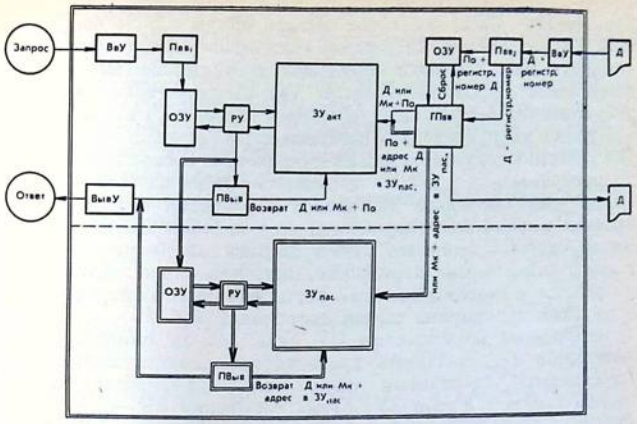



Рис. 54. Обобщенная структурная блок-схема документальной ИПС

ВвУ — вводное устройство; ВывУ — выводное устройство; ГПвв — главный преобразователь ввода (выполняет следующие функции: 1) преобразует документы в форму, пригодную для ввода в $ЗУ_{акт.}$ или $ЗУ_{пас.}$; 2) вырабатывает адреса документов в $ЗУ_{пас.}$; 3) записывает в $ЗУ_{акт.}$ поисковые образы документов с их микрокопиями или адресами этих документов в $ЗУ_{пас.}$); Д — документ; $ЗУ_{акт.}$ — активное запоминающее устройство; $ЗУ_{пас.}$ — пассивное запоминающее устройство; Мк — микрокопия документа; ОЗУ — оперативное запоминающее устройство; $Пвв_1$ — преобразователь ввода 1-й (производит перевод информационного запроса на ИПЯ, используемый в данной ИПС); $Пвв_2$ — преобразователь ввода 2-й (описывает основное смысловое содержание документа в терминах ИПЯ, используемого в данной ИПС); Пвыв — преобразователь вывода (преобразует результаты информационного поиска в требуемую форму); По — поисковый образ документа; РУ — решающее устройство. Двойными линиями выделены элементы второго контура ИПС, а двойными стрелками показаны функциональные связи между этими элементами, ввод в $ЗУ_{акт.}$ поисковых образов документов и адресов хранения этих документов (или их микрокопий), ввод в $ЗУ_{пас.}$ документов (или их микрокопий) с присвоенными им адресами, передача промежуточных результатов информационного поиска из первого контура во второй, а также передача конечных результатов информационного поиска из второго контура ИПС в ВывУ (при работе ИПС в двухконтурном варианте)

рис. 54 приведена обобщенная структурная блок-схема такой ИПС. При этом необходимо отметить, что органической частью такой ИПС является человек, без активного участия которого, по крайней мере в процессе индексирования документов и информационных запросов, нормальное функционирование системы пока невозможно.

Центральными узлами документальной ИПС являются активное и пассивное запоминающие устройства. Активное запоминающее устройство ($ЗУ_{акт.}$) предназначено для хранения поисковых образов документов с самими документами, их микрокопиями или адресами хранения этих документов (или их микрокопий) в пассивном запоминающем устройстве ($ЗУ_{пас.}$). Именно при помощи $ЗУ_{акт.}$ и соответствующего решающего устройства реализуется информационный поиск, т. е. производится сопоставление

хранящихся в $ЗУ_{акт.}$ поисковых образов документов с вводимыми в ИПС поисковыми предписаниями, проверяется выполнение установленного критерия смыслового соответствия между ними и принимается решение о выдаче или невыдаче того или иного документа. $ЗУ_{пас.}$, как видно из самого названия, никакой активной роли в информационном поиске не играет и используется лишь для хранения документов или их микрокопий. В связи с этим необходимо указать, что под ЗУ понимается любая материальная система, предназначенная для записи и хранения информации во времени с целью воспроизведения этой информации по запросу. Такой системой являются любой письменный документ, книга, ящик с каталожными карточками, книгохранилище, специальные устройства с магнитным, электростатическим, фотографическим и другими принципами записи информации и т. п.

Основными показателями ЗУ являются их емкость, время обращения и длительность хранения информации. Емкость ЗУ определяется предельным количеством информации, которое можно записать в этом ЗУ. Время обращения ЗУ — это время, которое нужно затратить на отыскание в ЗУ нужной ячейки и считывание (или запись) в ней информации. Время обращения ЗУ принято обозначать символом $t_{обр.}$. В некоторых современных ЗУ это время доведено до долей микросекунды.

ЗУ могут быть классифицированы по их функциональным свойствам, схемно-конструктивным принципам, физико-химическим принципам действия запоминающих элементов, по назначению ЗУ и т. п. В отношении возможностей обращения к некоторой ячейке ЗУ эти устройства могут быть разделены на следующие три группы:

а) ЗУ *с последовательным доступом*, в которых обращение к заданной ячейке требует как бы последовательного прохождения мимо всех других ячеек. Примером таких ЗУ являются устройства с магнитной лентой или рулонным микрофильмом;

б) ЗУ *с циклическим доступом*, в которых возможность обращения к той или иной ячейке периодически через некоторое время повторяется. Примером таких устройств являются ЗУ на магнитных барабанах и дисках;

в) ЗУ *с произвольным доступом*, которые в любой момент допускают обращение к любой из ячеек. Примером таких ЗУ являются устройства матричного типа [1].

ЗУ состоит из ячеек. Под ячейкой ЗУ в данном случае понимается устройство, предназначенное для хранения одного поискового образа документа вместе с самим документом, его микро-

[1] Подробные сведения о принципах действия, конструкциях и основных характеристиках ЗУ различных типов содержатся в кн. Л. П. Крайзмера «Устройства хранения дискретной информации» М.—Л., Госэнергоиздат, 1961.

копией или адресным шифром; одной лексической единицы дескрипторного ИПЯ вместе с адресными шифрами документов, в поисковые образы которых входит эта лексическая единица; одного адресного шифра документа вместе с самим этим документом или его микрокопией. Ячейкой ЗУ является каталожная карточка, унитерм-карта, перфокарта, диамикрокарта, кадр микрофильма, зона магнитной ленты и т. п.

Возможны четыре следующих варианта заполнения ячейки ЗУ в документальной ИПС:

$$1)\ P_i \to C_i,$$
$$2)\ P_i \to a_i,$$
$$3)\ d_i \to \langle a_{i1},\ a_{i2},\ a_{i3}\ldots,\ a_{i\gamma}\rangle,$$
$$4)\ a_i \to C_i,$$

где $P_i$ — поисковый образ документа; $C_i$ — документ или его микрокопия; $a_i$ — адрес хранения документа или его микрокопии; $d_i$ — лексическая единица дескрипторного ИПЯ; $a_{i1}$, $a_{i2}$, $a_{i3}$, . . ., $a_{i\gamma}$ — адреса хранения документов, поисковые образы которых содержат $d_i$.

В зависимости от того, находится поисковый образ документа на одном общем носителе с самим документом (или его микрокопией) или физически отделен от него, документальная ИПС имеет одно или два ЗУ. Если используется вариант $P_i \to C_i$ для заполнения ячеек ЗУ, то ИПС имеет только одно активное ЗУ.

Все остальные варианты заполнения ячеек требуют применения в ИПС двух ЗУ — активного и пассивного. Естественно задать вопрос, какой из двух возможных схем структурного построения ИПС следует оказать предпочтение — одноконтурной или двухконтурной.

Сопоставление характеристик действующих в настоящее время ИПС, рассчитанных на большие объемы документальной информации, приводит к выводу, что предпочтение должно быть оказано схеме с двумя ЗУ — активным и пассивным, так как эта схема имеет следующие основные преимущества перед схемой с одним активным ЗУ:

1. Более высокую разрешающую способность. Применение быстродействующей ЭЦВМ для выявления по заданным признакам адресов документов, содержащих требуемую информацию, позволяет более полно и без замедления процесса информационного поиска учитывать базисные и текстуальные отношения, в которые вступают дескрипторы поисковых образов документов.

Кроме того, раздельное хранение поисковых образов документов и самих документов (или их микрокопий) создает для человека возможность, если в том возникает необходимость, корректировать информационный поиск, просматривая и уточняя

промежуточные результаты поиска в $ЗУ_{акт.}$ перед их передачей в $ЗУ_{нас.}$.

2. Более высокое быстродействие, особенно при рациональной организации $ЗУ_{акт.}$. В двухконтурной ИПС информационный поиск производится в два этапа: сначала выявляются адреса документов с требуемой информацией, а затем по этим адресам в $ЗУ_{нас.}$ отыскиваются и сами документы или их микрокопии. Если для выявления адресов искомых документов применить быстродействующую ЭЦВМ (что и делается в некоторых больших ИПС), то этот этап информационного поиска выполняется очень быстро — иногда всего за несколько секунд. Когда выявлены адреса документов с требуемой информацией, то время обращения в $ЗУ_{нас.}$ на втором этапе информационного поиска также не может быть большим. В наиболее совершенных $ЗУ_{нас.}$ с произвольным доступом $t_{обр.}$ не превышает 5 сек. (вместе со временем изготовления копии документа). Таким образом, весь процесс поиска требуемой информации в массиве, содержащем сотни тысяч и миллионы страниц документов, может производиться в таких ИПС за 20—30 сек. машинного времени.

В одноконтурной ИПС для выполнения информационного поиска необходимо последовательно просмотреть адресные части ячеек всего ЗУ или определенной его секции. Ячейки ЗУ в данном случае выполняются в виде соответствующих зон рольного микрофильма или в виде унитарных дискретных носителей (диамикрокарт, видеомагнитных карт или апертурных перфокарт) и должны при поиске механически перемещаться относительно устройства считывания. Таким образом, в одноконтурных ИПС применяются ЗУ последовательного доступа с механической системой перемещения носителя информации. ЗУ такого типа имеют большое время обращения, изменяющееся прямо пропорционально их емкости.

3. Минимальный износ микрокопий документов в $ЗУ_{нас.}$. В ИПС первого типа носители информации из-за постоянных перемещений довольно быстро изнашиваются, и поэтому их нужно периодически заменять новыми. Например, в ИПС типа «Minicard» для этой цели хранится полный комплект негативов всех диамикрокарт, введенных в ее ЗУ. Эти негативы используются лишь для регенерации приходящих в негодность позитивных диамикрокарт, предназначенных непосредственно для информационного поиска.

4. Более широкие возможности для реорганизации поисковых образов документов.

5. Возможность строить такие ИПС из имеющегося оборудования. В зависимости от конкретных условий двухконтурные ИПС можно создавать, применяя в качестве $ЗУ_{акт.}$ соответствующие ЭЦВМ, счетно-перфорационные машины, комплекты суперпозиционных перфокарт и т. п., а в качестве $ЗУ_{нас.}$ — микро-

фильмовые селекторы, картотеки из форматных микрофильмов, диа- и эпимикрокарт стандартных форматов и т. п.

Важным вопросом, встающим перед конструкторами современных документальных ИПС, является отношение к опыту информационной работы, накопленной в научно-технических библиотеках. Если внимательно разобраться, то в общие принципы документального информационного поиска пока внесено сравнительно мало такого, что в том или ином виде не применялось бы ранее в библиотечной практике. Более того, самые совершенные из современных ИПС, в сущности, представляют собой лишь варианты автоматизированных библиотек, в которых все или почти все рабочие операции выполняются не людьми, а машинами. Ранее была приведена обобщенная структурная блок-схема современной документальной ИПС. При желании функционирование этой схемы можно описать и в операциях обычной библиотечно-библиографической работы. Такое описание будет иметь следующий вид:

1. Читатель обращается в информационно-библиографический отдел библиотеки с просьбой подобрать литературу по интересующему его вопросу. Его информационный запрос записывается, например, на специальном требовании. Эта операция эквивалентна подаче информационного запроса в ВвУ ИПС.

2. Библиограф переводит поступивший запрос на информационно-поисковый язык, используемый в данной библиотеке (например, представляет его в виде некоторой совокупности индексов УДК). В ИПС перевод запроса на информационный язык системы производится в $\text{Пвв}_1$.

3. Индексы УДК, обозначающие основную тему запрошенной информации, выписываются на требование, которое направляется в книгохранилище. Эта операция эквивалентна вводу поискового предписания, переведенного на информационно-поисковый язык ИПС, в ОЗУ системы.

4. Если литература в книгохранилище библиотеки расставлена по индексам УДК, то работник книгохранилища, получив требование с выписанными на нем индексами, отыскивает издания, на которых проставлены те же индексы УДК, что и в требовании. Отобранная литература передается на кафедру выдачи.

Если же в библиотеке применяется форматно-порядковая система расстановки литературы, то поиск сначала производится не в книгохранилище, а в соответствующем каталоге. В результате этого поиска устанавливаются шифры изданий, содержащих требуемую информацию. Эти шифры выписываются на требование, которое затем передается в книгохранилище для подбора литературы. В ИПС этому этапу работы библиотечного персонала соответствует поиск в $\text{ЗУ}_{\text{вкт}}$ документов или адресов документов, в которых содержится требуемая информация. Этот поиск производится по поисковым образам, введенным в ОЗУ. Соответствую-

щее устройство считывания просматривает адресные части ячеек ЗУ, и поисковые образы документов, считываемые этим устройством, сравниваются с поисковыми предписаниями. При совпадении РУ выдает команду на считывание информации, хранящейся в соответствующей ячейке ЗУ.

Если ячейка ЗУ выполнена в виде диамикрокарты, видеомагнитной карты или апертурной перфокарты, на которой вместе с поисковым образом документа приводится и его микрокопия, то считывание такой ячейки производится путем ее временного физического извлечения из ЗУ. Если же поисковые образы документов и сами документы (или их микрокопии) хранятся в ИПС раздельно (ситуация, которой соответствует форматно-порядковая расстановка литературы в книгохранилище библиотеки), то в результате поиска в $ЗУ_{акт.}$ выдаются не сами документы или их копии, а лишь адреса искомых документов. По этим адресам затем производится поиск самих документов $ЗУ_{пас.}$.

5. Отобранные издания, в которых содержится требуемая информация, используются для подготовки ответа на информационный запрос. В зависимости от требования могут быть изготовлены токопии или микрофильмы этих документов; может быть также тавлен библиографический список отобранных изданий, подовлены их аннотации или рефераты. В ИПС все эти операции оизводятся в Пвыв.

6. После подготовки ответа на запрос отобранные издания возвращаются в книгохранилище. То же происходит и в ИПС.

Из приведенного сравнения видно, что любая документальная ИПС по своей функциональной схеме пока мало чем отличается от обычной научно-технической библиотеки. И это не случайно. Проблема поиска научной информации возникла одновременно с возникновением науки, т. е. несколько тысяч лет назад. Правда, до последнего времени эта проблема по многим причинам не имела такой остроты, какую она приобрела в настоящее время. Но такая проблема существовала, и ее все это время удовлетворительно решали библиотеки. Разумеется, с течением времени эта задача усложнялась. Соответственно все более совершенствовались и методы информационного поиска, применявшиеся в библиотечной практике. Но все операции при информационном поиске по-прежнему выполнялись людьми.

В результате невиданно быстрого увеличения объема публикуемой в мире документальной информации, которым сопровождается развитие современной науки и техники, сложилась качественно новая ситуация. Человек в силу определенной ограниченности своих возможностей более не может, пользуясь лишь традиционными библиотечно-библиографическими методами и средствами, удовлетворительно решать задачу документального информационного поиска. Возникла острая практическая потребность в создании машинных ИПС, которые могли бы значительно

эффективней, чем человек, выполнять формально-логические операции информационного поиска.

За последние 15—20 лет был создан целый ряд машинных ИПС различных типов. Однако все эти ИПС пока относятся к классу систем «человек—машина», т. е. не могут функционировать без участия человека. Полностью автоматические ИПС могут быть созданы лишь тогда, когда будут формализованы все логические операции информационного поиска.

Рис. 55. Схема прямой организации $ЗУ_{акт.}$ ИПС

Проводя аналогию между современной ИПС и обычной научно-технической библиотекой, мы не стремились в какой-то мере умалить принципиальную новизну и сложность научных и технических задач, стоящих перед создателями современных ИПС, а лишь хотели обратить их внимание на то, что было бы неправильно пренебрегать ценным опытом библиотечно-библиографической работы, накопленным за тысячелетия существования библиотек.

Ранее были приведены четыре возможных варианта заполнения ячеек $ЗУ_{акт.}$ и $ЗУ_{пас.}$ в ИПС. Из них в $ЗУ_{акт.}$ могут быть использованы только первые три варианта, поскольку в четвертом варианте не содержится поискового образа документа. Варианты $P_i \rightarrow C_i$ и $P_i \rightarrow a_i$, в сущности, есть частные случаи одной и той же схемы. Отсюда следует, что вообще возможны лишь две схемы заполнения ячеек $ЗУ_{акт.}$, которые и будут рассмотрены ниже.

1. В адресную часть каждой ячейки $ЗУ_{акт.}$ ИПС записывается некоторый поисковый образ, а ее информационную часть составляет оригинал, микрокопия или адрес хранения документа, которому поставлен в однозначное соответствие этот поисковый образ. В принципе возможна ситуация, при которой в информационной части одной ячейки $ЗУ_{акт.}$ будут собраны все оригиналы, микрокопии или адреса хранения документов, имеющие идентич-

ные поисковые образы. Такая схема организации $ЗУ_{акт.}$ ИПС называется *прямой* (или *подокументной*).

На рис. 55 приведена матрица, наглядно раскрывающая принцип прямой организации $ЗУ_{акт.}$. На этой матрице каждая строка поставлена в однозначное соответствие номеру (адресу хранения) документа или его микрокопии, а каждый столбец — дескриптору

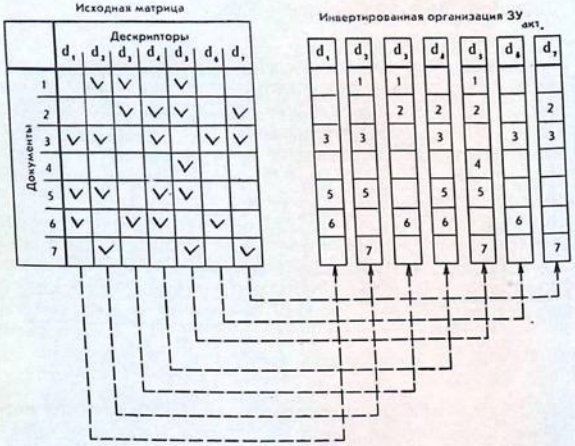



Рис. 56. Схема инвертированной организации $ЗУ_{акт.}$ ИПС

$d_i$, входящему в основной словарный состав ИПЯ. Если в поисковом образе документа содержится дескриптор $d_i$, то на пересечении строки с номером этого документа и столбца $d_i$ в матрице сделана отметка. Справа от матрицы показано, какую информацию содержат отдельные ячейки $ЗУ_{акт.}$ при такой схеме их заполнения. В правой (адресной) части каждой ячейки (строки) записан поисковый образ соответствующего документа.

При такой схеме организации $ЗУ_{акт.}$ процедура информационного поиска состоит из считывания информации, которая записана в адресной части каждой ячейки $ЗУ_{акт.}$, и сопоставления этой информации с поисковым предписанием в целях проверки соблюдения критерия их смыслового соответствия.

2. На рис. 56 показана другая схема организации $ЗУ_{акт.}$, которая называется *инвертированной* (ее можно назвать также *словарной*). При использовании такой схемы в адресной части каждой ячейки $ЗУ_{акт.}$ записывается некоторый дескриптор $d_i$, а в ее информационной части — номера (адреса хранения) всех

документов, в поисковые образы которых входит этот дескриптор. На рис. 56 справа от матрицы выделены ее столбцы, которые изображают ячейки $ЗУ_{акт.}$ ИПС. Верхние графы этих столбцов — суть адресные части ячеек.

При такой схеме заполнения ячеек $ЗУ_{акт.}$ процедура информационного поиска состоит из следующих двух этапов. Сначала в $ЗУ_{акт.}$ необходимо найти ячейки, которые поставлены в однозначное соответствие дескрипторам, образующим поисковое предписание, причем в информационных частях этих ячеек должны быть считаны все записанные там номера документов. Следующий этап процедуры — выявление таких номеров документов, которые одновременно содержатся во всех без исключения ячейках $ЗУ_{акт.}$, считанных на первом этапе поиска. Если номер документа содержится во всех таких ячейках, то это означает, что в поисковый образ документа с таким номером входят все дескрипторы, составляющие поисковое предписание, т. е. что данный документ является искомым.

Необходимо отметить, что рассмотренные процедуры столь просты лишь в тех случаях, когда в ИПС применяются дескрипторные языки без грамматики. Если же в ИПЯ применяется та или иная система грамматических средств для выражения текстуальных отношений между дескрипторами в поисковых образах документов и в поисковых предписаниях, то процедуры информационного поиска значительно усложняются. При использовании дескрипторного ИПЯ с грамматикой для решения вопроса о выдаче документа недостаточно установления факта вхождения всех дескрипторов поискового предписания в поисковый образ этого документа. Необходимо еще проверить, находятся ли дескрипторы поискового образа в таких же текстуальных отношениях между собой, как и в поисковом предписании. Применяемые для этого методы здесь не рассматриваются, так как это выходит далеко за рамки данной книги.

Ниже приводится описание основных методов реализации документального информационного поиска в элементарных терминах теории множеств. Пусть $A$ — некоторое множество документов, в котором производится информационный поиск, а $D$ — множество дескрипторов, образующих основной словарный состав ИПЯ

$$A = \{a_1,\ a_2,\ a_3,\ \ldots,\ a_n\},$$

где $a_1,\ a_2,\ a_3,\ \ldots,\ a_n$ — отдельные документы или их номера (адресные шифры);

$$D = \{d_1,\ d_2,\ d_3,\ \ldots,\ d_m\},$$

где $d_1, d_2, d_3\ \ldots, d_m$ — отдельные дескрипторы. Документальный информационный поиск состоит в выделении из множества $A$ такого подмножества $A$ $(d_i^*)$, каждому элементу которого (т. е. каж-

дому документу) поставлен в однозначное соответствие поисковый образ $P_i$. Обозначим множество всех $P_i$ через $P^*$:
$P_i = \{d_{i1}, d_{i2}, d_{i3}, \ldots, d_{i\alpha}\} = \bigcup\limits_{\alpha} d_{i\alpha}$, $1 \leqslant \alpha \leqslant m$ ($m$ — число элементов в $D$);

$$P_i = \subset P^*;$$

$$P^* = P_1 \cup P_2 \cup P_3 \cup \ldots \cup P_k = \bigcup\limits_{k} P_k = D, \quad 1 \leqslant k \leqslant m.$$

Пусть $P_q$ — поисковое предписание. В простейшем случае критерием отбора документов из собрания $A$ является вхождение в документ (в явном или скрытом виде) заданного множества дескрипторов $P_q = \bigcup\limits_{\alpha} d_{i\alpha}$. Документ считается отвечающим на информационный запрос $Q$ и отбирается из собрания документов $A$, если его поисковый образ $P_i$ эквивалентен $P_q$. Очевидно, что если не введено специальное ограничение, то информационному запросу $Q$ будут по смыслу соответствовать и такие документы, поисковые образы которых имеют вид:

$$P_i = \bigcup\limits_{\delta} d_{i\delta}, \quad \delta = \alpha + \rho, \quad \rho = 1, 2, 3, \ldots, k; \quad \alpha + \rho \leqslant m.$$

Это означает, что на $A$ должны быть также отобраны документы, для которых имеет место отношение $P_q \subset P_i$, т. е. когда $P_q$ — правильная часть $P_i$.

Возможны два основных метода решения этой задачи.

1. *Последовательный поиск.* $ЗУ_{акт.}$ ИПС должно содержать не менее $n$ ячеек ($n$ — число элементов в $A$). Заполнение каждой ячейки $ЗУ_{акт.}$ производится по следующей схеме:

$$P_i \rightarrow a_i,$$

где $P_i = \bigcup\limits_{\alpha} d_{i\alpha}$, $a_i \in A(d_i^*)$, $A(d_i^*) = \{a_{i1}, a_{i2}, a_{i3}, \ldots, a_{i\beta}\}$, $1 \leqslant \beta \leqslant n$.

Такая схема организации $ЗУ_{акт.}$ в документальной ИПС называется прямой[2]. В зависимости от средств технической реализации ИПС в качестве $a_i$ может выступать как сам документ (или его микрокопия), так и адрес (номер) этого документа в $ЗУ_{нак.}$.

1. 1. При использовании $ЗУ_{акт.}$ с прямой организацией эквивалентность $P_i$ и $P_q$ может быть установлена лишь путем сопоставления каждого $P_i$ и $P_q$. Это означает, что для реализации информационного поиска необходимо последовательно просмотреть адресные части всех ячеек $ЗУ_{акт.}$ ИПС (или определенной части этого $ЗУ_{акт.}$), так как поисковый образ каждого документа записан в адресной части ячейки. Поэтому такой метод документального информационного поиска называют последовательным поиском.

В $ЗУ_{акт.}$ с прямой организацией информационный поиск производится по четко заданным адресам, в качестве которых выступают

[2] Возможен также вариант $P_i \rightarrow A(d_i^*)$.

поисковые образы документов. Поэтому такой информационный поиск можно было бы также назвать адресным поиском.

Рассмотрим основные принципы сопоставления $P_i$ и $P_q$ при последовательном обращении в $ЗУ_{акт.}$ документальной ИПС.

А. $P_i$ и $P_q$ эквивалентны в том, и только в том, случае, если по некоторому правилу каждому элементу из $P_q$ можно поставить в соответствие один, и только один, элемент из $P_i$, так что каждый элемент из $P_i$ окажется поставленным в соответствие одному, и только одному элементу $P_q$. Отсюда следует, что (1—1)-соответствие между элементами конечных множеств $P_i$ и $P_q$ может быть установлено в том, и только в том, случае, когда число элементов в $P_i$ равно числу элементов в $P_q$, т. е. когда $P_i$ и $P_q$ — равночисленны.

Однозначное соответствие между элементами множеств $P_i$ и $P_q$ может быть установлено путем выполнения над этими множествами операций сложения, пересечения или вычитания. $P_i$ и $P_q$ эквивалентны тогда, и только тогда, когда имеет место любая из следующих четырех пар равенств:

$$\begin{array}{ll} 1)\ P_i \cup P_q = P_q, & P_i \cup P_q = P_i; \\ 2)\ P_i \cap P_q = P_q, & P_i \cap P_q = P_i; \\ 3)\ P_i / P_q = \varnothing, & P_q P_i = \varnothing; \\ 4)\ P_i \cup P'_q = P^*, & P_i \cap P'_q = \varnothing; \end{array} \qquad (1)$$

$$(P'_q = P^* P_q, \quad P_q \cap P'_q = \varnothing; \quad P_q \cup P'_q = P^*).$$

Б. Пусть $\alpha$ — число элементов в $P_q$, а $\delta$ — число элементов в $P_i$. При $\alpha < \delta$ соотношение $P_q \subset P_i$ будет иметь место в том случае, если можно установить (1—1)-соответствие между элементами $P_q$ и $P'_i \subset P_i$. Для доказательства соотношения $P_q \subset P_i$ необходимо получить в результате операций сложения или пересечения $P_i$ и $P_q$ и операции сложения $P_i$ и $P'_q$ любое из следующих равенств:

$$\begin{array}{l} 1)\ P_i \cup P_q = P_i, \\ 2)\ P_i \cap P_q = P_q, \\ 3)\ P_i \cup P'_q = P^*. \end{array} \qquad (2)$$

Последний способ доказательства соотношения $P_q \subset P_i$ чаще всего применяется тогда, когда ячейки $ЗУ_{акт.}$ ИПС выполнены в виде диамикрокарт или зон рольного микрофильма. Этот вариант был применен, например, в первых моделях ИПС типа «Rapid Selector», в сканирующем устройстве Г. Луна, в ИПС типа «Universal Card Scanner» и в ИПС типа «Filmorex».

Из сопоставления равенств (1) и (2) следует, что при последовательном информационном поиске лишь операции сложения или пересечения $P_i$ и $P_q$ и операции сложения $P_i$ и $P'_q$ позво-

ляют проверить, имеет ли место между этими множествами соотношение $P_q \subset P_i$ и возможно ли соотношение $P_i \equiv P_q$.

Очевидно, что описанный метод документального информационного поиска может быть применен как в одноконтурных, так и в двухконтурных ИПС, т. е. в ИПС с одним или двумя ЗУ — активным и пассивным.

1.2. Продолжительность последовательного информационного поиска $\tau$ по запросу $Q$ есть функция от $n$ ($n$ — число элементов в $A$). При использовании прямой схемы заполнения ячеек $ЗУ_{акт.}$ ИПС записи $P_i \to a_i$ обычно располагаются в $ЗУ_{акт.}$ в последовательности ввода документов в ИПС. Однако если $n$ в $A$ достаточно велико, то для уменьшения $\tau$ при определенных условиях может быть применен метод разбиения множества всех введенных в $ЗУ_{акт.}$ ИПС записей $P_i \to a_i$ на взаимоисключающие классы.

Пусть $M^*$ — множество всех введенных в $ЗУ_{акт.}$ ИПС записей вида $P_i \to a_i$. Множество $M^*$ может быть разбито на классы $M_1, M_2, M_3, \ldots, M_k$; $\bigcup\limits_k M_k = M^*$, $\bigcap\limits_k M_k = \varnothing$, $k \ll n$. Тогда информационный поиск по любому запросу $Q$ будет производиться не в $M^*$, а в одном из таких классов $M_i \subset M^*$. Очевидно, что число элементов в любом множестве $M_i$ будет меньше $n$. Если $e$ — число элементов в $M_i$, а $\tau_i$ — продолжительность информационного поиска по запросу $Q$ в $M_i$, то из $e < n$ следует, что $\tau_i < \tau$.

Однако применение такого метода сокращения $\tau$ ограничено тем, что в качестве основания для разбиения $M^*$ на взаимоисключающие классы можно использовать лишь такие характеристики, как год издания документа; страна, в которой документ опубликован; язык, на котором документ написан, и т. п., и очень трудно использовать для этой цели предметные характеристики документов.

Нетрудно понять, что разбиение $M^*$ на взаимоисключающие классы технически значительно легче производить в тех случаях, когда ячейки $ЗУ_{акт.}$ ИПС выполнены в виде унитарных дискретных носителей (перфокарт, апертурных перфокарт, диамикрокарт, магнитных и видеомагнитных карт).

Другой метод уменьшения продолжительности информационного поиска в ИПС с прямой организацией $ЗУ_{акт.}$ — разбиение $M^*$ (по любому основанию) на непересекающиеся подмножества $M_1, M_2, M_3, \ldots, M_h$ $\left(\bigcup\limits_k M_h = M^*,\ 1 < h \ll n\right)$ и поиск по одному и тому же запросу $Q$ одновременно во всех этих подмножествах. Тогда продолжительность информационного поиска по запросу $Q$ будет составлять $r/h$; так как $h > 1$, то $r/h < \tau$.

Аналогичный метод мультиплексии можно применить и для повышения производительности ИПС рассматриваемого типа,

осуществляя информационный поиск одновременно по нескольким запросам.

Очевидно, что применение метода мультиплексии для сокращения $\tau$ и повышения производительности ИПС при последовательном информационном поиске не требует какого-либо изменения принципов организации $ЗУ_{акт.}$. Для этого лишь необходимо, чтобы конструкция $ЗУ_{акт.}$ позволяла одновременно сканировать адресные части ячеек нескольких секций $ЗУ_{акт.}$ и сопоставлять записанные в этих ячейках $P_i$ с одним $P_q$ или в процессе каждого обращения в $ЗУ_{акт.}$ — с несколькими различными $P_q$.

Если в документальной ИПС допустимо существенное увеличение емкости применяемого $ЗУ_{акт.}$, то весьма эффективным методом сокращения $\tau$ будет группировка всех записей $P_i \to a_i$ по каждому входящему в $P_i$ дескриптору. В результате такой группировки мы получим $m$ классов записей $P_i \to a_i$ ($m$ — число элементов в $D$, т. е. число всех возможных дескрипторов). В $ЗУ_{акт.}$ ИПС эти классы обычно располагаются в определенном порядке (например, в алфавитном порядке дескрипторов, образующих классы записей $P_i \to a_i$). При такой организации $ЗУ_{акт.}$ ИПС для выполнения информационного поиска по запросу $Q$ не требуется последовательного просмотра адресных частей всех ячеек $ЗУ_{акт.}$. Для этого достаточно произвести такой просмотр лишь в одной секции $ЗУ_{акт.}$, поставленной в однозначное соответствие некоторому классу записей $P_i \to a_i$. Выбор секции $ЗУ_{акт.}$ для информационного поиска в данном случае производится по любой характеристике $d_i \in P_q$. Поскольку число записей вида $P_i \to a_i$ в каждой такой секции будет, как правило, значительно меньше $n$ ($n$ — число документов, введенных в ИПС), то на последовательный просмотр адресных частей всех ячеек в любой такой секции $ЗУ_{акт.}$ потребуется значительно меньше времени, чем на выполнение такой же операции во всем этом $ЗУ_{акт.}$.

1.3. Выше было указано, что выявление соотношений $P_i \equiv P_q$ и $P_q \subset P_i$ производится путем сопоставления элементов $P_i$ и $P_q$. Для этого достаточно лишь однократно сканировать адресные части ячеек $ЗУ_{акт.}$ без считывания записанной в этих ячейках информации. Такой метод последовательного информационного поиска, называемый *методом сканирования*, позволяет производить информационный поиск с достаточно большой скоростью. Если записи $P_i \to a_i$ располагаются в $ЗУ_{акт.}$ в порядке их ввода в ИПС, а устаревание документальной информации есть функция от продолжительности хранения этой информации в ИПС, то устаревшими будут считаться все документы, введенные в ИПС ранее некоторой контрольной даты. Такую документальную информацию можно легко выводить из ИПС, если установить корреляцию между номерами (адресами) документов и календарным временем их ввода в данную ИПС.

Однако существуют сложные информационно-поисковые задачи, которые трудно или даже невозможно решить путем однократного сканирования адресных частей ячеек $ЗУ_{акт.}$ с одновременной проверкой на совпадение $P_i$ и $P_q$. В таких задачах факт эквивалентности (или неэквивалентности) между $P_i$ и $P_q$ может быть установлен лишь в результате выполнения более или менее длинной цепочки взаимосвязанных логических операций, для чего необходимо многократное обращение в $ЗУ_{акт.}$ ИПС. Последовательный информационный поиск такого типа был назван поиском по методу итерационного анализа. Предназначенные для такого поиска ИПС характеризуются прерывистым действием с последовательным чередованием двух следующих действий:

а) перезаписи информации из определенной зоны $ЗУ_{акт.}$ ИПС в оперативное ЗУ;

б) итерационного анализа информации, введенной в оперативное ЗУ.

При определенных условиях записи в $ЗУ_{акт.}$ ИПС, предназначенной для поиска информации по методу итерационного анализа, могут быть короче записей, предназначенных для информационного поиска по методу непрерывного сканирования. Вообще существует много переменных, определяющих, который из этих двух методов последовательного информационного поиска — метод непрерывного сканирования или метод итерационного анализа — является оптимальным для каждой конкретной ситуации [3].

Подавляющее большинство информационно-поисковых задач, требующих применения итерационного анализа, относится к фактографическому информационному поиску, который в данном случае не рассматривается.

2. *Выборочно-сопоставительный поиск.* $ЗУ_{акт.}$ должно иметь не менее $m$ ячеек ($m$ — число элементов в $D$). Заполнение каждой ячейки $ЗУ_{акт.}$ производится по следующей схеме:

$$d_i \rightarrow A(d_i),$$

где $d_i \in P_i$, $A(d_i) \subset A$, $A(d_i) = \{a_{i1}, a_{i2}, a_{i3}, \ldots, a_{i\gamma}\}$, $1 \leqslant \gamma \leqslant n$.

В данном случае $a_i$ — адресный шифр (номер), который документ с дескриптором $d_i$ имеет в $ЗУ_{пас.}$ ИПС. Такая схема организации $ЗУ_{акт.}$ называется *инвертированной*. $ЗУ_{акт.}$ с инвертированной организацией обычно бывает упорядоченным по $\{d_i\}$. Если $\{d_i\}$ — дескрипторы, то в $ЗУ_{акт.}$ они располагаются в углубленном алфавитном порядке.

Схема $d_i \rightarrow A(d_i)$ требует физического отделения поисковых образов документов от самих документов. Это означает, что такая схема может быть применена лишь в ИПС с двумя ЗУ — активным и пассивным.

---

[3] Такая ситуация возникает, например, когда дескрипторы в $P_i$ и $P_q$ находятся в сложных текстуальных отношениях между собой.

Ввод записей $d_i \to a_i$ в $ЗУ_{акт.}$ может производиться как подокументно (т. е. по всем дескрипторам $\{d_i\}$, образующим $P_i$ одного данного документа), так и по группам документов (т. е. по всем характеристикам $\{d_i\}$, использованным в поисковых образах всех документов данной группы). Существует вероятность того, что некоторые дескрипторы из $\{d_i\}$ в поисковых образах различных документов данной группы будут повторяться. Поэтому метод ввода записей $d_i \to a_i$ в $ЗУ_{акт.}$ по группам документов (если это допускают конкретные условия) может оказаться более рациональным, чем метод подокументного ввода. Применение метода ввода записей $d_i \to a_i$ по группам документов позволяет избежать многократного отыскания в $ЗУ_{акт.}$ ячеек, поставленных в однозначное соответствие элементам из $\{d_i\}$, которые повторяются в поисковых образах определенной группы вводимых в ИПС документов.

Пусть дан поисковый образ $P_q = \bigcup_{\alpha} ad_{i\alpha}$. Требуется выбрать из $A$ такое подмножество $A(d_i^*)$, каждый элемент которого содержит непременно все дескрипторы $\{d_{i1}, d_{i2}, d_{i3}, \ldots, d_{i\alpha}\}$. При этом возможны два случая.

А. $\alpha = 1$. Для решения задачи необходимо и достаточно отыскать в упорядоченном $ЗУ_{акт.}$ ИПС ячейку $d_i$ и считать из нее $A(d_i)$, которое и есть искомое подмножество документов.

Б. $\alpha > 1$. Тогда решение задачи производится в два этапа:

1) путем сопоставления с элементами $P_q$ дескрипторов $\{d_i\}$, внесенных в адресные части ячеек $ЗУ_{акт.}$, отыскиваем все ячейки $ЗУ_{акт.}$, которые поставлены в однозначное соответствие дескрипторам $d_i \in P_q$, и из этих ячеек выписываем всю содержащуюся в них информацию. В результате выполнения такой операции получаем следующую таблицу:

$$
\begin{aligned}
&d_{i1} \to A(d_{i1}),\\
&d_{i2} \to A(d_{i2}),\\
&d_{i3} \to A(d_{i3}),\\
&\ldots\ldots\\
&d_{i\alpha} \to A(d_{i\alpha});
\end{aligned}
$$

2) отобранные дескрипторы $\{d_i\}$ образуют множество $P_i$, которое эквивалентно $P_q$ (поскольку в результате произведенной операции между элементами $P_i$ и $P_q$ было установлено (1—1)-соответствие). Теперь необходимо из множеств $A(d_{i1})$, $A(d_{i2})$, $A(d_{i3})$, $\ldots$, $A(d_{i\alpha})$ выделить такое подмножество $A(d_i^*)$, каждый элемент которого непременно содержит все дескрипторы $\{d_{i1}, d_{i2}, d_{i3}, \ldots, d_{i\alpha}\}$.

Очевидно, что все дескрипторы $\{d_{i1}, d_{i2}, d_{i3}, \ldots, d_{i\alpha}\}$ одновременно содержатся лишь в тех элементах $A(d_{i1})$, $A(d_{i\alpha})$, $A(d_{i3})$, $\ldots$,

$A(d_{ia})$, которые входят в каждое из этих множеств, образуя их пересечение $A(d_i^*)$:

$$A(d_i^*) = A(d_{i1}) \cap A(d_{i2}) \cap A(d_{i3}) \cap \ldots \cap A(d_{ia}).$$

Для того чтобы ускорить реализацию операции пересечения множеств $\{A(d_i)\}$, каждое множество $A(d_i)$ должно быть упорядочено по адресам (номерам) входящих в него документов.

После нахождения $A(d_i^*)$ из $ЗУ_{нас.}$ ИПС выбираются документы, отвечающие информационному запросу $Q$. Задача решена.

Итак, в $ЗУ_{акт.}$ с инвертированной организацией при $\alpha > 1$ (типичный случай) выявление адресов (номеров) искомых документов производится в два этапа. Сначала в $ЗУ_{акт.}$ разыскиваются и считываются ячейки, отвечающие всем входящим в них дескрипторам. Розыск этих ячеек производится по дескрипторам, каждый из которых выполняет функцию координаты некоторого подмножества документов. Затем в считанных ячейках путем сопоставления выявляются номера документов, одновременно содержащиеся во всех этих ячейках. Эти номера — суть адресные шифры искомых документов. Такой метод документального информационного поиска может быть назван *выборочно-сопоставительным* поиском (ср. с англ. look-up and compare — выбрать и сопоставить). Как было указано выше, выборочно-сопоставительный метод поиска может быть применен лишь в ИПС с двумя ЗУ — активным и пассивным.

В ИПС с инвертированной организацией $ЗУ_{акт.}$ документы располагаются в порядке их ввода в систему. Недостатком схемы $d_i \rightarrow A(d_i)$ является то, что для вывода из ИПС устаревших документов необходимо производить перезапись адресов документов во всех ячейках $ЗУ_{акт.}$ (если устаревание документальной информации считать функцией от продолжительности ее хранения в ИПС). От этого недостатка можно избавиться, если применить схему $d_i \rightarrow a_i$ и выполнять ячейки $ЗУ_{акт.}$ в виде унитарных дискретных носителей (например, перфокарт машинной сортировки).

Сравнение прямой и инвертированной схем организации $ЗУ_{акт.}$ ИПС показывает, что инвертированная схема имеет три основных недостатка:

1. Необходимость соблюдения фиксированной последовательности расположения ячеек в $ЗУ_{акт.}$, когда в качестве таких ячеек используются перфокарты.

2. Трудность выражения и учета при информационном поиске текстуальных отношений между дескрипторами в поисковых образах документов.

3. Более сложную процедуру ввода поисковых предписаний в ИПС.

Кроме того, машинные ИПС с инвертированной организацией $ЗУ_{акт.}$ начали создаваться значительно позже, чем ИПС с прямой организацией $ЗУ_{акт.}$. Поэтому в настоящее время ИПС с прямой организацией $ЗУ_{акт.}$ часто обладают более высокими эксплуатационными характеристиками (большим быстродействием, более высокой точностью выдач и меньшей стоимостью информационного поиска), чем ИПС с инвертированной организацией $ЗУ_{акт.}$. Это особенно справедливо в случаях применения ИПС с прямой организацией $ЗУ_{акт.}$ для решения таких сложных информационно-поисковых задач, как поиск химических соединений по фрагментам их структурных формул и т. п.

С другой стороны, ИПС с инвертированной организацией $ЗУ_{акт.}$ обладают и определенными преимуществами перед ИПС с прямой организацией $ЗУ_{акт.}$.

Если применяется ИПЯ без грамматики, то ИПС с инвертированной организацией $ЗУ_{акт.}$ в принципе должны иметь большее быстродействие, чем ИПС с прямой организацией $ЗУ_{акт.}$, особенно при информационном поиске в сравнительно больших массивах документов.

Бесспорным преимуществом инвертированной организации $ЗУ_{акт.}$ перед прямой является также то, что при определенных условиях она позволяет путем введения обратной связи корректировать процесс поиска по промежуточным результатам. Информационный поиск с такой обратной связью очень напоминает известную игру в вопросы и ответы. В этой игре участвуют два человека. Игра состоит в том, что один из партнеров задумывает, например, фамилию какого-либо ученого, писателя, поэта, общественного деятеля и т. п., а другой путем постановки некоторого числа вопросов, на которые он может получить лишь ответы «да» или «нет», должен угадать задуманную фамилию. Разумеется, эта фамилия должна быть известна угадывающему. По условиям игры число вопросов обычно ограничивается 15—20. Опыт показывает, что путем постановки достаточно продуманной последовательности все более уточняющих вопросов можно успешно решать даже очень сложные задачи такого типа.

Аналогичную стратегию поиска можно применить и при инвертированной организации $ЗУ_{акт.}$ ИПС. Для этого лишь необходимо, чтобы оператор ИПС имел возможность в процессе такого информационного поиска обращаться в целях контроля непосредственно к текстам документов, выделяемых из поискового массива в ответ на ввод в ИПС очередного дескриптора из числа дескрипторов поискового предписания. Такой метод последовательных приближений применен в ИПС типа «Microcite», в которой ячейки $ЗУ_{акт.}$ выполнены в виде суперпозиционных перфокарт.

При прямой организации $ЗУ_{акт.}$ применение обратной связи для корректировки процесса поиска по промежуточным результатам связано с большими трудностями.

Опыт показывает, что преимущества инвертированной схемы организации $ЗУ_{акт.}$ перед прямой увеличиваются по мере того, как уменьшается отношение $\frac{m}{n}$ ($m$ — число дескрипторов в ИПЯ без грамматики, $n$ — число документов, введенных в ИПС). При $\frac{m}{n} \geqslant 1$ более предпочтительным становится прямая организация $ЗУ_{акт.}$.

Очевидно, что отношение $\frac{m}{n}$ будет тем меньше, чем больше $n$. Но с увеличением $n$ растет и $m$. Зависимость $m$ от $n$ может быть приблизительно описана следующими эмпирическими формулами [4]:

$$m = 4170 \log_{10}(n + 730) - 11620$$

и

$$\frac{dm}{dn} = \frac{1810}{+730}.$$

Из этих формул видно, что с увеличением $n$ скорость роста $m$ быстро уменьшается. Отсюда можно сделать вывод, что если применяется дескрипторный ИПЯ без грамматики, то с увеличением числа введенных в ИПС документов инвертированная схема организации $ЗУ_{акт.}$, казалось бы, становится все более предпочтительной перед прямой.

Но это не так. С ростом массива введенных в ИПС документов возникает необходимость в повышении точности выдач этой ИПС. Когда объем поискового массива достигнет некоторой критической величины, потребуется введение в дескрипторный ИПЯ соответствующих грамматических средств. При таких условиях применение в ИПС инвертированной организации $ЗУ_{акт.}$ существенно затруднит информационный поиск, сделает его очень дорогостоящим.

Отсюда можно сделать общий вывод, что инвертированная организация $ЗУ_{акт.}$ ИПС безусловно предпочтительнее прямой, когда в ИПС применяется дескрипторный ИПЯ без грамматики или с элементами грамматики, т. е. когда поисковый массив не очень велик. Если же применяется дескрипторный ИПЯ с достаточно развитой системой грамматических средств, то в отношении экономической эффективности (т. е. стоимости и времени информационного поиска) прямая схема организации $ЗУ_{акт.}$ ИПС может быть значительно выгодней, чем инвертированная.

В заключение данного раздела следует отметить, что при построении документальной ИПС выбор схемы организации $ЗУ_{акт.}$

---

[4] E. Wall. Information systems. «Chemical Engineering Progress», 1955, v. 55, № 1, стр. 55—59.

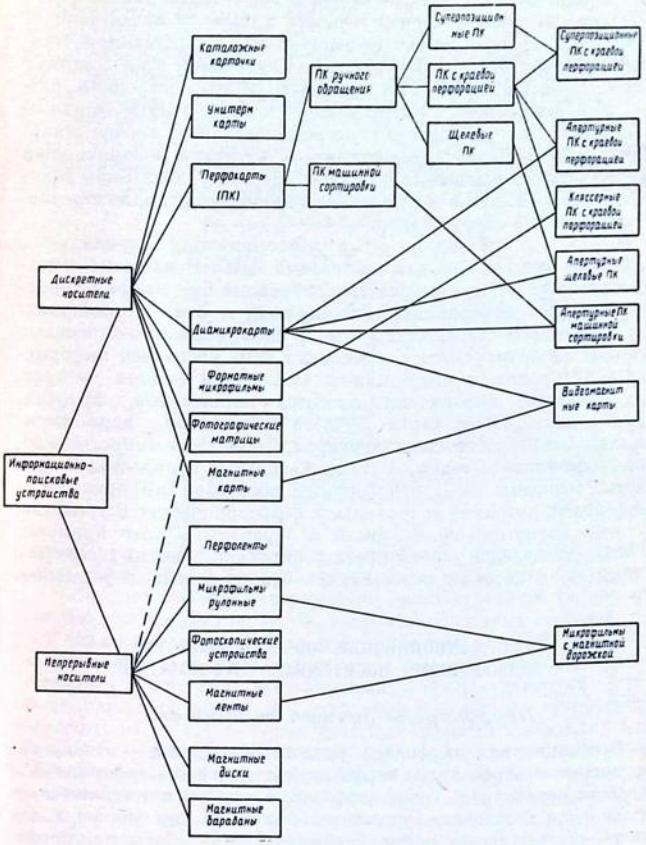



Рис. 57. Схема классификации информационно-поисковых устройств (по характеру носителя информации)

Суперпозиционные перфокарты были изобретены еще в 1915 г. американским орнитологом Г. Тейлором, предложившим использовать такие перфокарты для идентификации птиц по различным сочетаниям их характерных признаков. Однако для документального информационного поиска суперпозиционные перфокарты начали применяться лишь в 40-х годах нашего столетия. Суперпозиционные перфокарты широко известны также под названием *просветные перфокарты*. В Великобритании их называют еще *характеристическими картами* (feature cards), а в США — *перфокартами типа «пик-э-бу»* (*амер.* pick-a-boo — вид игры в прятки, распространенный в США).

Выше было указано, что запись адреса документа на суперпозиционной перфокарте производится путем пробивки перфокарты в точке, координаты которой поставлены в однозначное соответствие адресу этого документа в $ЗУ_{нас.}$ ИПС. Очевидно, что на перфокарте можно сделать ограниченное число пробивок. Это число прямо пропорционально полезной площади перфокарты и плотности размещения пробивок на перфокарте и обратно пропорционально площади каждой пробивки. Предельное число пробивок, которое можно сделать на суперпозиционной перфокарте, и составляет информационную емкость этой перфокарты, т. е. предельное число адресов документов, которое вмещает данная перфокарта. Ограниченная информационная емкость является, по-видимому, основным недостатком суперпозиционных перфокарт. Повышение плотности размещения пробивок на суперпозиционной перфокарте и уменьшение площади каждой пробивки существенно усложняют перфорацию и считывание. Поэтому основным методом повышения информационной емкости суперпозиционных перфокарт является увеличение их линейных размеров. В настоящее время выпускаются суперпозиционные перфокарты размером от 187×82,5 *мм* до 445×445 *мм* и емкостью от 400 до 40 тыс. документов. Минимальный диаметр пробивки на таких перфокартах — 0,6 *мм*, максимальный — 3 *мм*.

Подавляющее большинство суперпозиционных перфокарт имеет координатную сетку, которая значительно облегчает процедуру пробивки перфокарт и определение номеров, поставленных в (1—1)-соответствие отверстиям на перфокарте.

На суперпозиционных перфокартах, не имеющих координатной сетки, определение номеров совпадающих пробивок производится при помощи трафарета из прозрачного материала, на который нанесена такая координатная сетка. Для пробивки суперпозиционных перфокарт обычно применяются специальные перфораторы вырубного и сверлильного типа, а для выявления пробивок, совпадающих в наборе таких перфокарт, используются подсвечиваемые экраны.

Суперпозиционные перфокарты хранятся в картотеках, где они расстанавливаются в углубленном алфавитном порядке де-

скрипторов, обозначающих эти перфокарты. Иногда суперпозиционные перфокарты снабжаются краевой перфорацией, которая позволяет быстро находить перфокарты с нужными дескрипторами и в неупорядоченном массиве.

Суперпозиционные перфокарты являются простым и вместе с тем весьма эффективным средством механизации документального информационного поиска в массивах, содержащих до 100 тыс. документов. Однако они применимы для построения лишь таких ИПС, в которых применяются дескрипторные языки без грамматики.

В 1961 г. в Национальном бюро стандартов США создана ИПС типа «Microcite», в которой $ЗУ_{акт.}$ выполнено в виде комплекта суперпозиционных перфокарт размером 203×127 *мм* и емкостью 18 тыс. документов [6]. Перфокарты изготовлены из поливинилового листа толщиной 0,25 *мм*. $ЗУ_{нас.}$ этой ИПС представляет собой горизонтально расположенный барабан из прозрачного материала, на котором закреплена съемная матрица размером 610×508 *мм*. Эта матрица содержит 18 тыс. микроизображений рефератов, отпечатанных на карточках формата 127×76 *мм*. Кратность уменьшения — 37,75 : 1. Над поверхностью подсвечиваемого экрана, который предназначен для выявления совпадающих пробивок в наборах суперпозиционных перфокарт, установлены подвижные координатные линейки $x$ и $y$. Эти линейки механически связаны с матричным барабаном. Перемещение $x$-линейки сопровождается соответствующим сдвигом матричного барабана вдоль его оси, а $y$-линейки — поворотом этого барабана вокруг оси. Перемещение координатных линеек и матричного барабана производится путем вращения двух приводных барабанов. Общая схема устройства ИПС типа «Microcite» показана на рис. 59.

Информационный поиск при помощи ИПС типа «Microcite» производится следующим образом. На подсвечиваемый экран ИПС укладывается набор суперпозиционных перфокарт, отвечающих дескрипторам поискового предписания. Пробивки, совпадающие во всех этих перфокартах, обозначают адреса искомых документов. Если необходимо ознакомиться с рефератом какого-либо из найденных документов, то для этого достаточно установить координатные линейки таким образом, чтобы обозначающая адрес этого документа пробивка оказалась на пересечении координатных линеек. Тогда изображение реферата интересующего документа проецируется в натуральную величину на выводной экран, который расположен на передней панели ИПС. Предусмотрена

---

[6] A new Microcite machine for large scale information searching. «NBS Technical News Bull.», 1961, v. 45, № 7, стр. 113—115; T. C. Bagg, M. E. Stevens. Information selection systems retrieving replica copies: a state-of-the art report. Washington, U. S. Government Printing Office, 1961, стр. 83—86 (NBS Technical Note 157).

возможность автоматического изготовления копии найденного реферата на рулонной фотобумаге, а также вывод адресов хранения документов в цифровой форме (на перфоленте или в печатном виде на бумажной ленте). Среднее время поиска первого документа (с выводом изображения его реферата на экран) — 1 мин., а каждого следующего документа — около 3 сек.

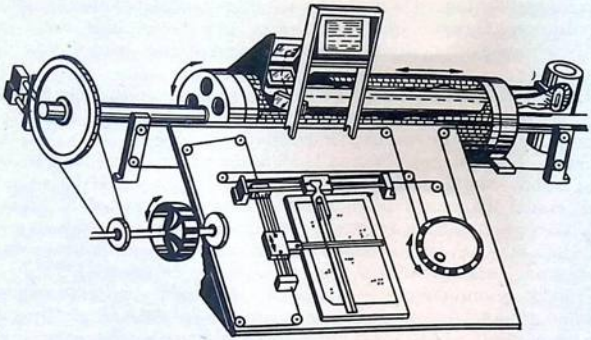


Рис. 59. Схема устройства ИПС типа «Microcite»

Перфокарты с краевой перфорацией[7]. Перфокарты с краевой перфорацией предназначены для построения документальных ИПС с прямой организацией $ЗУ_{акт.}$. Хотя первый патент на перфокарту с краевой перфорацией был выдан еще в 1896 г., практическое применение перфокарты такого типа получили лишь с 1920 г. Перфокарта с краевой перфорацией представляет собой прямоугольник из плотной бумаги, вдоль краев которого нанесен один или несколько рядов калиброванных отверстий — *краевая перфорация*. Каждое такое отверстие образует двоичную ячейку, в которой может быть записана «1» или «0», т. е. символы наличия или отсутствия какой-либо характеристики. Запись «1» в той или иной ячейке производится путем вырезки перемычки, отделяющей отверстие от края перфокарты, что превращает это отверстие в открытую щель.

Для отбора перфокарт, которые имеют вырезку в определенной ячейке (т. е. на каждой из которых записана соответствующая характеристика), достаточно ввести сортировальную спицу в пачки

[7] H. F. McGraw. Marginal punched cards in college and research libraries. Washington. The Scarecrow Press, 1952; M. Scheele. Punched-card methods in research and documentation. With special reference to biology. New York, Interscience Publishers, 1961.

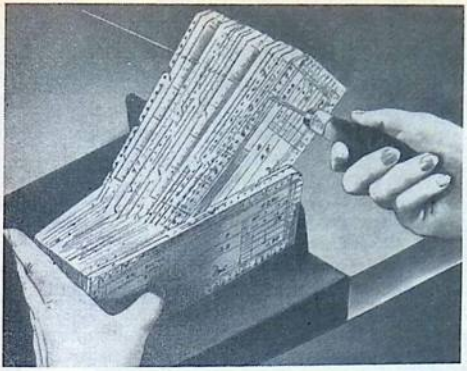


Рис. 60. Принцип использования перфокарт с краевой перфорацией

перфокарт и слегка встряхнуть эту пачку. Тогда из нее выпадут все перфокарты, имеющие вырезку в данной ячейке. Конструкция и принцип использования перфокарт с краевой перфорацией показаны на рис. 60.

Карточки типа «Zato» (см. рис. 63) не имеют предварительно нанесенной краевой перфорации. Вместо нее верхний край такой карточки снабжен специальной разметкой, которая показана на рисунке. Запись характеристики на карточке типа «Zato» производится путем вырезки открытой щели в соответствующем пункте верхнего края этой карточки. Для отбора карточек типа «Zato», имеющих вырезки в определенных позициях, пачка таких карточек (200 шт.) укладывается в специальный вибрационный селектор. В этом селекторе сортировальные спицы устанавливаются в позициях, соответствующих позициям краевых вырезок в искомых карточках. Пачка сортируемых карточек укладывается на эти спицы той стороной, на которой сделаны краевые вырезки. Тогда при встряхивании пачки вниз сдвигаются лишь те карточки, которые имеют вырезки в позициях, совпадающих с позициями сортировальных спиц. Опустившиеся карточки можно легко отделить от остальных, если пропустить сортировальный стержень в специальное отверстие, расположенное у нижнего края каждой карточки (в селекторе такое отверстие будет находиться в верхней части пачки карточек, так как они укладываются в селектор в опрокинутом положении). Выпускается два варианта карточек типа «Zato»: а) с разметкой только вдоль верхнего края (40 ячеек); б) с разметкой вдоль верхнего и нижнего краев (72 ячейки).

в однозначное соответствие опередеденная конфигурация кодовых символов, записываемая в некотором участке (субполе) кодового поля, отведенном для данной группы характеристик. Каждое кодовое слово, записанное в селекторном коде, должно иметь одно и то же число кодовых символов. В одном субполе кодового поля можно записать по селекторному коду только одну характеристику из данной группы характеристик. В этом состоит основной недостаток селекторного кода, который удобен лишь для записи взаимоисключающих характеристик. Число различных характеристик $C$, которое можно записать по селекторному коду в некотором поле, имеющем $n$ ячеек, если кодовое слово состоит из $m$ символов, можно определить по формуле:

$$C=\frac{n!}{m!\,(n-m)!}.$$

Значение $C$ максимально, когда $m=\frac{1}{2}n$. Если $m=1$, то мы имеем прямой код, который можно рассматривать как предельный случай селекторного кода.

При записи цифр от 0 до 9 по селекторному коду каждая цифра представляется в субполе из 6 ячеек в виде комбинации (суммы) двух элементов ряда X, 0, 1, 2, 4, 7. Каждому элементу этого ряда приведена в (1—1)-соответствие определенная ячейка кодового поля. В цифровом селекторном коде цифры от 0 до 9 представляются следующими комбинациями элементов X, 0, 1, 2, 4, 7:

$$0=X+0 \qquad 5=1+4$$
$$1=X+1 \qquad 6=2+4$$
$$2=X+2 \qquad 7=X+7$$
$$3=1+2 \qquad 8=1+7$$
$$4=X+4 \qquad 9=2+7$$

На рис. 61, *б* показана запись в селекторном коде цифры 9 на перфокарте с одним рядом краевой перфорации. Следует отметить, что в практических условиях цифра 0 часто записывается в кодовом поле не в виде комбинации (X+0), а по прямому коду.

При использовании перфокарт с двумя рядами краевой перфорации для записи цифр от 1 до 9 может быть применен вариант селекторного кода, при котором цифрам 1, 2, 4 и 7 приведены в (1—1)-соответствие пары перфораций (отверстий). Кодовое поле в таком случае должно состоять из $4\times2=8$ ячеек. Тогда каждая цифра ряда 1, 2, 4, 7 записывается на перфокарте одной глубокой вырезкой, а каждая цифра ряда 3, 5, 6, 8, 9 представляется в виде суммы соответствующей пары цифр ряда 1, 2, 4, 7 и записывается на перфокарте двумя мелкими вырезками. Число цифр, записанных в соответствующей группе кодовых полей

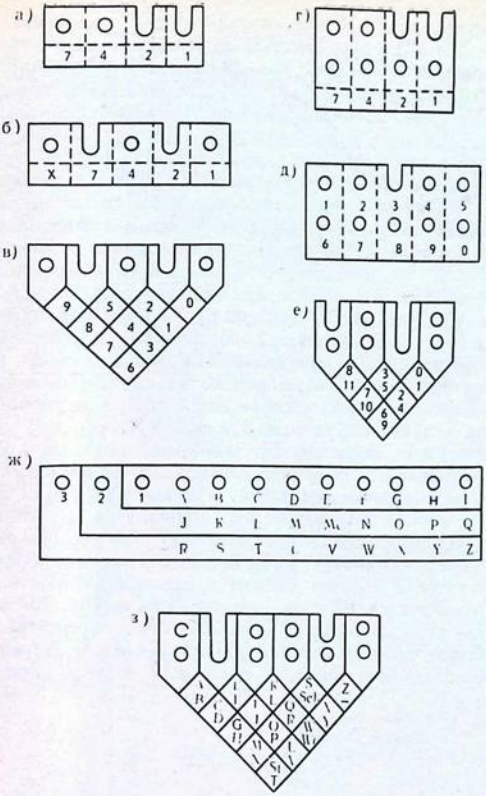


Рис. 61. Основные виды кодов для перфокарт с краевой перфорацией

перфокарты с двумя рядами краевой перфорации, указывается прямым кодом во вспомогательном поле, которое предшествует основным кодовым полям.

Частным случаем цифрового селекторного кода является *треугольный* или *пирамидальный цифровой селекторный код*, пример использования которого показан на рис. 61, *в*. В таком коде каждая цифра от 0 до 9 записывается двумя вырезками в ячейках, находящихся у верхних концов двух диагональных цифровых

рядов кодовой сетки, на пересечении которых находится записываемая цифра. Преимуществом треугольного селекторного кода перед линейным является то, что при его использовании для записи любой цифры ряда $1 \div 9$ требуется кодовое поле, содержащее не 6, а 5 ячеек. Очевидно, что треугольный селекторный код может применяться и для записи характеристик на перфокартах, имеющих два ряда краевой перфорации. На рис. 61, *е* показан пример записи в таком коде цифры 7. Глубокая вырезка справа означает, что из двух цифр, которые указаны на пересечении диагональных цифровых рядов с вырезками, записана верхняя цифра, т. е. цифра 7. Если необходимо записать в этом кодовом поле цифру 10, то глубокой должна быть левая вырезка. Таково общее правило записи цифр и букв в треугольном селекторном коде на перфокартах с двумя рядами краевой перфорации. Иногда цифры и буквы в квадратах, образуемых пересекающимися диагональными рядами треугольной кодовой сетки, располагаются не одна над другой, а рядом. Тогда глубокой вырезкой справа обозначается правая цифра или буква, а глубокой вырезкой слева — левая цифра или буква.

*Порядковым кодом* называется такой код, который предназначен для расположения перфокарт в нужном порядке за минимальное число сортировочных операций. При использовании порядкового кода каждой характеристике приводится в однозначное соответствие определенная конфигурация кодовых символов внутри некоторого кодового поля, причем эта конфигурация может состоять из любого числа кодовых символов, не превышающего числа ячеек в кодовом поле. Из этого определения следует, что двоичный код можно рассматривать как частный случай порядкового кода. Максимальное число различных характеристик, которые можно записать порядковым кодом в некотором поле, определяется по формуле:

$$C = \sum_{m=0}^{m=n} \frac{n!}{m!\,(n-m)!},$$

где $n$ — число ячеек в кодовом поле и $m$ — число кодовых символов в одном кодовом слове.

Преимущество порядкового кода перед другими комбинационными кодами состоит в том, что при его использовании для записи $k$ характеристик требуется значительно меньшее кодовое поле, чем при использовании прямого и селекторного кодов. Недостаток порядкового кода — возможность записи в каждом кодовом поле только одной характеристики из данной группы характеристик. Поэтому порядковый код особенно удобен для записи взаимоисключающих характеристик. В качестве примера порядкового кода можно назвать широко известный код 1-2-4-7

(так называемый *аддитивный код*). На перфокарте с одним рядом краевой перфорации в кодовом поле, состоящем из четырех ячеек (пробивок), можно записать по такому коду одной или двумя вырезками все цифры от 0 до 9 (см. рис. 61, *а*). Недостатком кода 1-2-4-7 является то, что он дает определенный «поисковый шум». Действительно, при отборе перфокарт с записью цифры 2 из массива будут выпадать также перфокарты, на которых записаны цифры 3 и 6, так как эти цифры составлены из сумм (1+2) и (2+4). Код 1-2-4-7 может быть применен и для перфокарт с двумя рядами краевой перфорации. Для записи любой цифры ряда от 0 до 9 в коде 1-2-4-7 на перфокарте с двумя рядами краевой перфорации требуется поле, состоящее из 4 пар ячеек. Если обозначить мелкие вырезки в таком поле символами $1'$, $2'$, $4'$, $7'$, а глубокие вырезки — символами $1''$, $2''$, $4''$, $7''$, то цифры ряда 0÷9 будут записываться следующим образом:

$$0 = 4' + 7' \qquad 5 = 1' + 4'$$
$$1 = 1'' \qquad 6 = 2' + 4'$$
$$2 = 2'' \qquad 7 = 7''$$
$$3 = 1' + 2' \qquad 8 = 1' + 7'$$
$$4 = 4'' \qquad 9 = 2' + 7'$$

На рис. 61, *г* показана запись 3 в коде 1-2-4-7 на перфокарте с двумя рядами краевой перфорации.

Выше уже было указано, что в зависимости от типа записываемых символов коды могут быть цифровыми и алфавитными. Для записи букв может быть применен прямой, селекторный или порядковый код. В принципе алфавитный код можно считать вариантом цифрового кода, так как каждой букве алфавита явно или в скрытом виде приводится в однозначное соответствие ее порядковый номер в этом алфавите. Если буквы алфавита записаны в порядковом коде, то это позволяет путем сортировки перфокарт быстро раскладывать их в алфавитном порядке. Примеры алфавитного кода приведены на рис. 61, *ж* и 61, *з*.

*Суперпозиционным кодом* называется такой код, который предназначен для записи в одном и том же кодовом поле нескольких характеристик. При использовании суперпозиционного кода каждой характеристике приводится в однозначное соответствие определенная конфигурация кодовых символов, причем число таких символов в каждом кодовом слове должно быть постоянным. Наиболее широко применяются следующие два варианта суперпозиционного кода:

а) характеристики записываются в одном общем кодовом поле (см. рис. 62 и 63);

б) характеристики записываются в кодовом поле, составленном из $M$ идентичных субполей, каждое из которых имеет $n$

ячеек. В таких субполях обычно записывается по одному кодовому символу.

Достоинством суперпозиционного кода является то, что он существенно повышает степень использования кодового поля перфокарты, что равносильно увеличению его емкости. Недостаток суперпозиционного кода — возможность появления «поискового шума» в результате образования ложных («паразитных») конфигураций при наложении одного на другой (суперпозиции) истинных кодов. Например, если в одном и том же кодовом поле записаны коды 17 и 52, то возникают ложные конфигурации 12, 15, 72 и 75. Некоторые из них могут совпасть с конфигурациями истинных кодов, что приведет к появлению «поискового шума». Чем больше кодов записывается в одном и том же поле, тем больше они накладываются один на другой и тем выше вероятность появления паразитных конфигураций. Когда применяется суперпозиционное кодирование, то при информационном поиске выдается некоторый статистически определимый процент документов, не имеющих никакого отношения к информационному запросу. Использование случайных чисел для построения кодовых слов при суперпозиционном кодировании снижает вероятность появления паразитных кодов. Для вычисления «поискового шума», возникающего при суперпозиционном кодировании, может быть использована следующая приближенная формула:

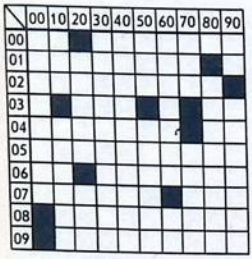



Рис. 62. Общий принцип суперпозиционного кодирования

$$F_d \approx (G/H)^Y,$$

где $F_d$ — величина «поискового шума» (т. е. доля выданных, но не относящихся к этому информационному запросу документов); $G$ — среднее число ячеек, занятых в суперпозиционном кодовом поле; $H$ — общее число ячеек в кодовом поле; $Y$ — число символов (элементов) в одном коде.

Отношение $G/H$ — правильная дробь ($G$ и $H$ — целые положительные числа и $G < H$; $0 < G/H < 1$). Отсюда следует, что $F_d$ тем меньше, чем больше $Y$. Если $G/H = {}^1/_2$, то при $Y=5$ $F_d = {}^1/_{32}$, а при $Y=12$ $F_d = {}^1/_{4096}$. Таким образом, при суперпозиционном кодировании «поисковый шум» тем меньше, чем больше кодовых символов содержит код. В практических условиях вели-

Рис. 63. Пример записи в суперпозиционном коде на карте типа «Zato»

чина $F_d$ для различных значений $H$ и $Y$ определяется по заранее вычисленным таблицам.

Таковы основные типы кодов, применяемых для записи характеристик в кодовых полях перфокарт с краевой перфорацией. В сущности, такие же коды используются и при работе со щелевыми перфокартами и перфокартами машинной сортировки.

В заключение раздела о перфокартах с краевой перфорацией необходимо указать, что перфокарты такого типа целесообразно использовать для механизации информационного поиска лишь в массивах до 10 тыс. документов — при ручной сортировке и до 30 тыс. документов — при использовании вибрационных или ударно-динамических селекторов. Но наиболее широко перфокарты с краевой перфорацией применяются для информационного поиска в массивах порядка одной-двух тысяч документов, а также для создания индивидуальных перфокартотек по узким вопросам науки и техники.

*Щелевые перфокарты* [8]. Как и перфокарты с краевой перфорацией, щелевые перфокарты предназначены для создания ИПС с прямой организацией $ЗУ_{акт.}$ Щелевая перфокарта, в сущности, представляет собой вариант перфокарты с краевой перфорацией. Кодовое поле щелевой перфокарты образовано прямоугольной матрицей из калиброванных отверстий. Расстояние между колонками и строками этой матрицы — постоянное.

Ячейкой кодового поля щелевой перфокарты является участок, расположенный между двумя смежными отверстиями (по вертикали или горизонтали, в зависимости от вида перфокарты). Запись характеристики в этой ячейке производится путем вырезки этого участка перфокарты, в результате чего на ее месте образуется щель.

Выпускаются щелевые перфокарты разных размеров и с разной емкостью кодового поля. Для записи характеристик на щелевых перфокартах применяются в основном такие же виды кодов, что и при работе с перфокартами с краевой перфорацией. В свободное поле щелевой перфокарты может быть занесено библиографическое описание документа, его аннотация или даже микрокопия, которая вклеивается в соответствующую апертуру (в виде диамикрокарты). Апертурные щелевые перфокарты выпускает, в частности, фирма Selection (Франция).

Поиск щелевых перфокарт, на которых записана совокупность тех или иных характеристик, производится следующим образом. Неупорядоченный массив перфокарт укладывается в сортировальный ящик специального селектора. В отверстия сортировального ящика и перфокарт, ограничивающие сверху (или сбоку) определенные ячейки кодового поля щелевой перфокарты, вставляются сортировальные спицы. После этого сортировальный ящик селектора опрокидывается на 180° (или 90°), и все перфокарты, в которых сделаны щелевые вырезки в этих ячейках, выдвигаются из массива вниз на длину щели. Когда сортировальный ящик селектора находится в опрокинутом положении, в специальное отверстие, расположенное у верхнего (или бокового) края перфокарт и сортировального ящика, вводится вспомогательная спица, которая связывает с ящиком основной массив перфокарт, но не связывает с ним выдвинувшиеся перфокарты (их края находятся ниже отверстия, через которое проходит данная спица). Затем ящик селектора приводится в нормальное положение; однако вспомогательная спица не дает выдвинувшимся перфокартам опуститься в массив остальных перфокарт. В таком положении из ящика селектора вынимаются основные сортировальные

---

[8] M. Scheele. Punched-card methods in research and documentation. With special reference to biology. New York, Interscience Publishers, 1961, стр. 67—86; Manuel pratique de reproduction documentaire et de selection. (Publication FID № 353). Paris, Gauthier-Villars, 1964, стр. 228—237.

спицы, что позволяет извлечь из массива все перфокарты, в кодовом поле каждой из которых записана совокупность определенных характеристик (например, поисковый образ документа).

Основное достоинство щелевых перфокарт состоит в том, что они лучше, чем перфокарты с краевой перфорацией, приспособлены для поиска документов одновременно по нескольким характеристикам (т. е. для так называемого многоаспектного поиска). Однако отбор щелевых перфокарт по заданным характеристикам возможен только при помощи специального селектора и в этом — главный недостаток перфокарт такого типа.

### *Реализация документальных ИПС при помощи счетно-перфорационных машин*

Применение счетно-перфорационных машин для реализации документальных ИПС имеет исключительно большое практическое значение. Такие ИПС могут быть созданы везде, где имеются машиносчетные станции. В связи с этим здесь уместно привести следующие цифры: в 1962 г. в народном хозяйстве нашей страны использовалось свыше 4 тыс. комплектов счетно-перфорационных машин; в течение семилетки в СССР создано свыше 1000 новых машиносчетных станций.

Но сравнительно большая доступность — не единственное достоинство документальных ИПС, создаваемых на базе счетно-перфорационных машин. Такие ИПС пригодны для работы со сравнительно большими массивами документов (до 200 тыс. и более); они отличаются достаточно высокой эффективностью, а также просты и надежны в эксплуатации. При использовании упорядоченных (подсортированных) массивов перфокарт счетно-перфорационные машины можно успешно применять для информационного поиска в фондах, содержащих миллионы документов.

Носителями информации в ИПС, реализованных при помощи счетно-перфорационных машин, обычно являются 65-, 80- и 90-колонные перфокарты. Возможно также использование для этой цели апертурных перфокарт, о чем подробней будет сказано ниже. В таких ИПС может применяться как прямая (подокументная), так и инвертированная схема организации $ЗУ_{акт.}$ ИПС первого типа создаются главным образом на базе сортировальных машин (стандартных или модифицированных), а ИПС второго типа — на базе картоподборочных машин.

При прямой организации $ЗУ_{акт.}$ на каждой перфокарте в кодированном виде записывается поисковый образ документа и его адресный шифр. В некоторых ИПС на каждой перфокарте приводится также библиографическое описание документа и даже его аннотация. Если в качестве унитарного дискретного носителя информации в ИПС применяется апертурная перфокарта, то

изменения системы кодирования этот словарь можно расширить до 10 млн. слов.

Коды дескрипторов поискового образа записываются на перфокарте в произвольном порядке. Из практики известно, что при координатном индексировании поисковый образ любого документа обычно содержит не более 10—15 дескрипторов. Поэтому емкость кодового поля рассматриваемой перфокарты (24 дескриптора) вполне достаточна для записи поисковых образов подавляющего большинства документов. Если же поисковый образ документа содержит более 24 дескрипторов, то на такой документ заводится вторая перфокарта и т. д.

От сортировальной машины D3 ИПС отличается лишь тем, что может за один прогон отбирать из массива все перфокарты, содержащие код заданного дескриптора в любом из 24 субполей правой зоны перфокарты. Это достигается путем подключения к сортировальной машине D3 консольной приставки, содержащей небольшое запоминающее устройство и соответствующие логические схемы. Приставка подключается к сортировальной машине D3 при помощи кабеля. Когда приставка отключена, сортировальная машина может использоваться для выполнения обычных для таких машин операций.

Информационный поиск производится последовательно по каждому дескриптору поискового предписания. Коды этих дескрипторов вводятся в ИПС по одному при помощи цифровой клавиатуры, расположенной на верхней панели консольной приставки. На поиск по первому дескриптору поискового предписания в массиве из 10 тыс. перфокарт затрачивается 15 мин., по второму дескриптору — 3 мин., по третьему — 1 мин. и т. д. В качестве носителей информации в ИПС типа D3D могут быть применены и апертурные перфокарты. Естественно, что при этом сокращается емкость кодового поля перфокарт.

*ИПС американской фирмы Abbott Laboratories* предназначена для поиска документов (рефератов журнальных статей) по медицине, фармакологии, фармацевтической химии, стоматологии и ветеринарии [12]. В качестве носителей информации в ней применяются 80-колонные перфокарты, обрабатываемые на электронной статистической машине IBM-101 (скорость работы 450 *ПК/мин*). На 80-колонной перфокарте в суперпозиционном коде записывается поисковый образ документа и его адресный шифр (прямая организация $ЗУ_{акт}$). Для координатного индексирования документов в этой ИПС применяется дескрипторный ИПЯ, содержащий около 3 тыс. дескрипторов. Словарь дескрипторов выполнен в виде печатного списка и в виде алфавитной картотеки, составленной из 80-колонных перфокарт. Текстуальные отношения между

[12] W. A. Southern. Mechanized processing and retrieval of biomedical information. «Methods of Information in Medicine», 1962, v. 1, № 1, стр. 16—22.

дескрипторами в поисковых образах не выражаются. Каждый дескриптор закодирован 12-значным случайным числом. Например, дескриптор «пенициллин» кодируется случайным числом 063-123-264-304, дескриптор «эмульсия» — случайным числом 007-020-025-171 и т. д.

Для записи поискового образа документа на перфокарте отведены колонки 1—40. Ряды 0—9 этого поля (400 ячеек) используются для записи кодов дескрипторов. Код 063-123-264-304 записывается в этом поле следующим образом:

063 — пробивка позиции 3 в колонке 6,
123 — пробивка позиции 3 в колонке 12,
264 — пробивка позиции 4 в колонке 26,
304 — пробивка позиции 4 в колонке 30.

Расчет показывает, что в поле из 400 ячеек в суперпозиционном коде можно записывать не более 50—70 12-значных случайных чисел. Ряды 11 и 12 этого поля (всего 80 ячеек) отведены для записи в прямом коде 80 наболее часто встречающихся дескрипторов, что существенно расширяет возможности индексирования. Колонки 41—45 перфокарт отводятся для записи порядкового номера какого-либо одного дескриптора, а колонки 46—72 — для записи этого дескриптора в алфавитном коде. Эти колонки используются лишь во вспомогательных перфокартах, предназначенных для механизации процесса подготовки основной перфокарты. В колонках 73—74 записывается год публикации индексируемого документа, а в колонках 75—79 — адресный шифр его реферата. Колонка 80 используется для управления работой машины.

Ежегодно в ИПС вводится около 7500 статей. Статистическая машина IBM-101 за один прогон производит отбор перфокарт по четырем дескрипторам. На коммутацию поискового предписания затрачивается 4 мин. Время информационного поиска в массиве из 10 тыс. документов составляет 23 мин. В дальнейшем машину IBM-101 намечалось заменить более быстродействующей машиной IBM-108, обрабатывающей перфокарты со скоростью 1000 *шт/мин*. Основными недостатками описанной ИПС являются: а) недостаточность индексирования химических соединений по их номенклатурным названиям, а не по фрагментам структурных формул этих соединений; б) невозможность выражения текстуальных отношений между дескрипторами в поисковых образах документов.

*UCS*. Заслуживает упоминания также *ИПС типа «Universal Card Scanner»* (сокращенно UCS) [13], которая была создана фирмой International Business Machines Corp. в 1958 г. [14]

[13] Universal Card Scanner. — «Универсальное устройство для сканирования перфокарт».

[14] H. P. Luhn. The IBM Universal Card Scanner for punched card information searching systems. В кн.: «Emerging solutions for mechanizing the

дескрипторов в поисковом предписании, и являются искомыми документами данной системы.

Необходимо отметить, что если $ЗУ_{акт.}$ ИПС на перфокартах машинной сортировки организовано по инвертированной схеме и для реализации этой ИПС используется стандартная картоподборочная машина, то каждая перфокарта должна содержать только один адресный шифр документа. Это означает многократное увеличение поискового массива перфокарт, что фактически сводит к нулю основное преимущество инвертированной схемы организации $ЗУ_{акт.}$ перед прямой (подокументной) — значительно меньшее время информационного поиска. Обоснованность такого утверждения можно показать на следующем примере.

Пусть дан массив из 100 тыс. документов. Для координатного индексирования этих документов используется словарь из 3 тыс. дескрипторов. Средний объем поискового образа одного документа 12 дескрипторов. При организации $ЗУ_{акт.}$ ИПС по инвертированной схеме необходимо создать поисковый массив на 1,2 млн. перфокарт; этот массив должен быть разбит на 3 тыс. групп (по числу дескрипторов), в каждой из которых будет содержаться в среднем по 400 дескрипторов. Отсюда следует, что на информационный поиск по сочетанию двух дескрипторов в массиве из 100 тыс. документов необходимо около 2 мин. работы стандартной картоподборочной машины. Для решения такой же задачи при помощи ИПС типа D3D, которая была описана выше, потребуется почти 3 часа машинного времени.

Приведенные цифры, казалось бы, позволяют сделать вывод о безусловной предпочтительности применения в ИПС на перфокартах машинной сортировки инвертированной схемы организации $ЗУ_{акт.}$ и реализации таких ИПС при помощи стандартных картоподборочных машин. Однако этот вывод нуждается в весьма серьезной оговорке. Дело в том, что такой вывод справедлив лишь в тех случаях, когда ИПС типа D3D производит отбор перфокарт в неупорядоченном массиве, т. е. путем последовательного перебора всех 100 тыс. перфокарт. Если же этот массив будет упорядочен по каждому дескриптору, входящему в поисковые образы документов, то он также возрастет до 1,2 млн. перфокарт и будет разбит на те же 3 тыс. групп, содержащих в среднем по 400 перфокарт в группе. Тогда ИПС типа D3D будет решать эту же задачу не за 3 часа, как при неупорядоченном массиве перфокарт, а менее чем за 1 мин.

В 1957 г. американская фирма Documentation, Inc. предложила вариант ИПС, в котором фактически устранялся основной недостаток применения стандартных картоподборочных машин в ИПС с инвертированной организацией $ЗУ_{акт.}$ — необходимость в многократном увеличении поискового массива перфокарт.

Эта ИПС получила название COMAC [18]. Основная идея данного предложения заключается в том, что на каждой перфокарте записывается не один, а 37 адресных шифров. Это существенно сокращает объем поискового массива. Кроме того, промежуточные результаты сравнений записываются на вспомогательных перфокартах. Для выполнения этих операций стандартная картоподборочная машина должна быть соответствующим образом модифицирована [19].

Предложение фирмы Documentation, Inc. было реализовано в ИПС типа IBM-9900 «Special Index Analyzer» [20], созданной фирмой International Business Machines Corp. в 1958 г. В качестве основного носителя информации в этой ИПС используются 80-колонные перфокарты. На каждой перфокарте записывается до 12 шестиразрядных адресных шифров документов. При информационном поиске первый из сравниваемых субмассивов перфокарт переписывается на перфоленту, выполняющую функции промежуточного ЗУ. С этой перфолентой сравнивается второй субмассив перфокарт, и результаты сравнения также записываются на перфоленте. Затем с этими результатами сравнивается третий субмассив перфокарт и т. д. При информационном поиске ИПС типа IBM-9900 может выполнять и определенные логические операции. Например, она может быть настроена таким образом, чтобы выдавались адресные шифры лишь тех документов, в поисковые образы которых входят дескрипторы $A$ и $B$, но не входит дескриптор $C$. Конечные результаты информационного поиска автоматически воспроизводятся с перфоленты электрической пишущей машиной [21].

В заключение данного раздела необходимо остановиться на применении апертурных перфокарт, которые в последнее время получили очень широкое распространение. Апертурные перфокарты впервые начали применяться для хранения и поиска микрокопий документов (фотоснимков) Управлением стратегических служб США — предшественником сегодняшнего Центрального разведывательного управления — еще во время второй мировой войны. В настоящее время выпускается много различных типов и видов апертурных перфокарт. За рубежом их называют также перфокартами типа «Filmsort». Это название дала им фирма

---

[18] Comac—Continuous Multiple Access Collator — «Картоподборочная машина с непрерывным многоканальным доступом».

[19] M. Taube. The Comac: An efficient punched card collating system for the storage and retrieval of information. В кн.: «Emerging solutions for mechanizing the storage and retrieval of information». Washington, Documentation, 1959, стр. 72—85. (Studies in coordinate indexing, v. 5).

[20] Special Index Analyzer — «Специальный анализатор указателя».

[21] R. W. Murphy. The IBM 9900 Special Index Analyzer. В кн.: «Emerging solutions for mechanizing the storage and retrieval of information». Washington, 1959, стр. 86—111.

Рис. 64. Апертурная перфокарта машинной сортировки

Film 'N File, Inc. (США) [22], которая первой начала серийное производство таких перфокарт.

Основным типом апертурной перфокарты сейчас можно считать 80-колонную перфокарту с апертурой (калиброванным окном) формата D (30 × 41 *мм*), в которую вклеивается один кадр неперфорированного 35-миллиметрового микрофильма (рис. 64). На перфокарте апертура занимает поле от колонки 54 до колонки 76. Монтаж микрокадров в бланки апертурных перфокарт производится с рулонного микрофильма. Для этого разработано и серийно выпускается соответствующее оборудование — ручные устройства (производительность 400 кадров/час), полуавтоматические устройства (с такой же производительностью), а также автоматические установки (производительность 2 тыс. кадров/час). Создана также различная аппаратура для контактного микрокопирования с апертурных перфокарт на другие апертурные перфокарты, в которые предварительно вклеена диазопленка; читально-копировальные аппараты, позволяющие за 8—10 сек. изготовлять с апертурных перфокарт увеличенные копии документов (возможные форматы копий 22×28; 28×35; 30×46 и 46×61 *см*); автоматическая установка для микрофотографирования документов на пленку, вмонтированную в апертурные перфокарты, и проявления экспонированной пленки (производительность до 80 апертурных перфокарт/час) и другое оборудование.

Основное достоинство апертурных перфокарт состоит в том, что такие перфокарты позволяют сократить на 80% и более

[22] В настоящее время фирма Film 'N File, Inc. носит название Filmsort Co. и является отделением американской фирмы Miehle-Gross-Dexter, Inc.

площадь помещений, используемых для хранения документов в натуральном виде, а также обеспечивают быстрый поиск этих документов и изготовление их копий.

Быстрое развитие науки и техники позволяет создавать все более сложные машины. Но изготовление любой машины производится по чертежам, число которых растет пропорционально сложности машины. Достаточно указать, что для изготовления только планера реактивного бомбардировщика типа B-47 необходимо более 28 тыс. чертежей [23]. Другой пример. Американский реактивный самолет типа «Boeing 707» имеет грузоподъемность 26 $m$, а вес чертежей, которые необходимы для его изготовления, — около 23 $m$ [24]. В литературе указывается, что Министерство обороны США ежегодно расходует на изготовление и копирование чертежей свыше 2 млрд. долл. [25] Таких примеров имеется множество, и все они свидетельствуют об исключительно большой остроте проблемы хранения, поиска и размножения чертежей и другой технической документации. Одним из наиболее перспективных путей решения этой проблемы является использование стандартизированных апертурных перфокарт. Такая тенденция особенно четко проявляется в США.

Весьма ценным свойством апертурных перфокарт является то, что их обработку можно производить на стандартных сортировальных машинах. Однако при такой обработке некоторые апертурные перфокарты приходят в негодность уже после ста прогонов через машину. Это происходит главным образом потому, что щетки воспринимающего механизма сортировальной машины царапают поверхность фотопленки, вмонтированной в апертурную перфокарту. Такой недостаток можно устранить путем удаления из воспринимающего механизма сортировальной машины соответствующих щеток (если машина имеет 80 щеток) или путем замены электромеханического воспринимающего механизма фотоэлектрическим. Кроме того, в настоящее время разрабатываются варианты апертурных перфокарт, в которых микроизображение защищено от царапин специальной пленкой или вмонтировано в апертуру таким образом, что его поверхность находится ниже поверхности перфокарты.

Некоторые потребители считают, что апертурные перфокарты более целесообразно использовать для организации $ЗУ_{пас.}$ ИПС. В таких ИПС информационный поиск по смысловому содержанию производится на ЭЦВМ или счетно-перфорационных машинах с обычными (неапертурными) перфокартами; эти ИПС выдают адресные шифры искомых документов, по которым далее из строго

[23] C. P. Bourne. Methods of information handling. New York, Wiley, 1963, стр. 219.

[24] N. Stahl. Cost of engineering paper. «Product Engineering», 1962, v. 33, № 24, стр. 79—90.

[25] C. P. Bourne. Указ. соч., стр. 219.

упорядоченной картотеки вручную извлекаются апертурные перфокарты с микрокопиями. Именно такую структурную схему имеет ИПС типа «Intellofax», которая была создана в Центральном разведывательном управлении США еще в 1947 г. и до последнего времени считалась одной из основных ИПС этого учреждения. $ЗУ_{нет.}$ ИПС типа «Intellofax» выполнено на обычных 80-колонных перфокартах, общий фонд которых в 1960 г. составлял более 40 млн. шт., а $ЗУ_{нас.}$ — на форматных микрофильмах (до 1954 г.). С 1954 г. для этой цели начали применяться 80-колонные апертурные перфокарты. К январю 1961 г. в этой ИПС хранилось уже свыше 2,5 млн. таких апертурных перфокарт [26].

Общий вывод, который можно сделать на основании изложенного выше материала, состоит в том, что апертурные перфокарты, являясь весьма перспективным видом унитарных дискретных носителей информации, особенно удобны для хранения, поиска и распространения чертежей и других технических документов.

### *Информационно-поисковые устройства на диамикрокартах*

*Filmorex* [27]. Первая модель ИПС «Filmorex» была создана французским врачом Ж. Самэном в 1950 г. С тех пор в нее было внесено много существенных изменений и усовершенствований.

Носителем информации в ИПС «Filmorex» является диамикрокарта формата 35×60 *мм* (рис. 65). Диамикрокарта [28] разде-

[26] «Documentation, indexing, and retrieval of scientific information. A study of Federal and non-Federal science information processing and retrieval programs. 86th Congress, U. S. Senate, Committee on Government Operations, Document No. 113». Washington, U. S. Government Printing Office, 1961, стр. 16—17, 63—64; P. W. Howerton. A unitized microfilm record system: a case study. «Proceedings of the 11th Annual Meeting and Convention of the National Microfilm Association», Annapolis, National Microfilm Association, 1962, стр. 21—25.

[27] J. Samain. Filmorex. Une nouvelle technique de classement et selection des documents et des information. Paris, Tari, 1952; J. Samain. Documentation by the Filmorex technique. В кн.: «Shera J. H., Kent A., Perry J. W. (eds.) «Information systems in documentation». New York, Interscience Publishers, 1957, стр. 478—487; J. Samain. Eine neue Technik elektronischer Recherche für Information und Dokumentation. «Nachrichten für Dokumentation», 1958, Bd. 9, № 1, стр. 35—40; J. Descans. Le Filmorex. Resultats d'une annee d'exploitation. «Bull. des bibliotheques de France», 1959, v. 4, № 3, стр. 135—139; M. G. Dupoux. Selection de fiches documentaires emploi de l'appareil «Filmorex». «Mémorial de l'artillerie française», 1959, v. 33, № 3, стр. 535—553; J. Samain. La diffusion et la recherche des informations. Pour une organisation rationelle de la documentation. [Conference faite a la reunion de l'ICSU — AB, 11 juillet 1960], Paris, 1960; J. Samain. Classification and information retrieval. «Proceedings of ADIA-Conference. Frankfurt/Main vom 9—12 Juni 1959». Frankfurt/Main, Deutsche Gesellschaft für Dokumentation, 1961, стр. 273—285. (Beiheft zu dem Nachrichten für Dokumentation, № 8.

[28] В первой модели ИПС «Filmorex» применялись диамикрокарты формата 45×70 *мм*.

Рис. 65. Диамикрокарта ИПС «Filmorex»

лена на зоны $A$ и $B$. Зона $A$ используется для одного кадра с микроизображением документа, а зона $B$ — для кода поискового образа этого документа. Максимальная емкость кодового поля — 20 дескрипторов (400 двоичных единиц), но обычно в нем записывается не более 15 дескрипторов (300 двоичных единиц). Код каждого дескриптора занимает одну поперечную колонку в зоне $A$.

ИПС «Filmorex» состоит из следующих установок и устройств: съемочной камеры, гильотины для разрезки микрофильма на диамикрокарты, селектора, увеличительно-печатающей установки и рабочих картотек — предметной, вспомогательной и авторской. Съемочная камера ИПС «Filmorex» (рис. 66) — вертикального типа; она имеет два объектива — вертикальный и горизонтальный, а также два объектных стола. Вертикальный объектив используется для фотографирования кода поискового образа документа (кратность уменьшения 2,5 : 1), а горизонтальный — для фотографирования (через систему из двух зеркал) реферата или первой страницы документа (кратность уменьшения 5,9 : 1). Фотографирование документа и его поискового образа производится одновременно. В съемочной камере применяется рулонный неперфорированный микрофильм шириной 35 *мм*. Микрофильм имеет утолщенную подложку и противоскручивающий слой. Емкость бобины с микрофильмом 30 *м*. В 1963 г. фирма Filmorex s. a. сообщила о выпуске съемочной камеры типа 6045, которая позволяет фотографировать документы с уменьшением от 5 : 1 до 20 : 1 (в прежних моделях съемочной камеры кратность умень-

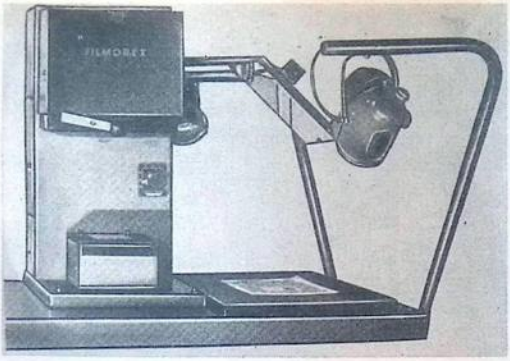



Рис. 66. Съемочная камера ИПС «Filmorex»

шения документов была фиксированной). Кроме того, при использовании этой камеры зона *A* диамикрокарты (кадр с микрокопией документа) может иметь длину 35, 40 и 45 *мм*, тогда как раньше длина зоны *A* также была фиксирована. В съемочной камере нового типа предусмотрена возможность изменения шага автоматической протяжки микрофильма. Этот шаг может иметь длину 20, 40 или 60 *мм*, что позволяет использовать съемочную камеру типа 6045 в качестве микрофильмирующей установки с размером кадров 20 × 32 и 40 × 32 *мм*. Модификация этой камеры — тип 642 — рассчитана на микрофильмирование документов размером до 80 × 120 *см* при кратности уменьшения от 4 : 1 до 30 : 1; она может быть использована для изготовления форматных микрофильмов разной длины.

В ИПС «Filmorex» для занесения на диамикрокарты поисковых образов документов применяются специальные кодовые карточки (рис. 67). Такие карточки [29] заводятся на каждый дескриптор и содержат 5-значный цифровой, буквенно-цифровой или буквенный шифр дескриптора, а также двоичный код этого шифра. Эти кодовые карточки используются следующим образом. Допустим, что поисковый образ документа состоит из дескрипторов *A*, *B*, *C* и *D*. Тогда оператор съемочной камеры извлекает из картотеки кодовые карточки этих дескрипторов и накладывает их одну на другую таким образом, чтобы каждая последующая

[29] В первых моделях съемочной камеры коды поисковых образов высвечивались на специальном экране (их ввод производился при помощи клавиатуры); для этой цели применялись также специальные перфокарты из черной бумаги, которые при фотографировании подсвечивались снизу и воспроизводили нужную кодовую мозаику.

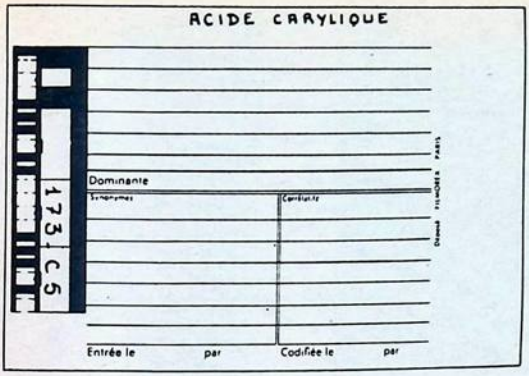



Рис. 67. Кодовая карточка с дескриптором Acide carylique для ИПС «Filmorex»

карточка закрывала все поле предыдущей, оставляя открытым лишь край с двоичным кодом дескриптора. В результате такого наложения кодовых карточек формируется код поискового образа документа, фотографируемый одновременно с документом. Последняя, верхняя кодовая карточка остается открытой. Поэтому вместе с кодом дескриптора фотографируется и его цифровой, буквенно-цифровой или буквенный шифр. На диамикрокарте этот шифр воспроизводится перед кодовым полем и легко читается невооруженным глазом. Основные достоинства такого метода формирования кодов поисковых образов — его простота и практически полное исключение возможности ошибок.

При вводе документа в ИПС «Filmorex» он фотографируется столько раз, сколько дескрипторов содержится в его поисковом образе, причем каждый раз последняя, верхняя кодовая карточка перекладывается вниз и становится первой. В результате этого открывается шифр следующего дескриптора, и этот шифр фотографируется на диамикрокарту. Таким образом, каждая последующая диамикрокарта оказывается обозначенной шифром очередного дескриптора. По этим шифрам диамикрокарты вручную раскладываются по соответствующим секциям рабочих картотек. Среднее время ввода документа в ИПС «Filmorex» около 2 мин. (при поисковом образе, состоящем из 5 дескрипторов).

Экспонированный рулонный микрофильм после проявления и технического контроля при помощи ручной гильотины разрезается на диамикрокарты. Производительность гильотины около 17 ДМК/мин. Готовые диамикрокарты раскладываются по секциям соответствующих картотек, где они хранятся в металли-

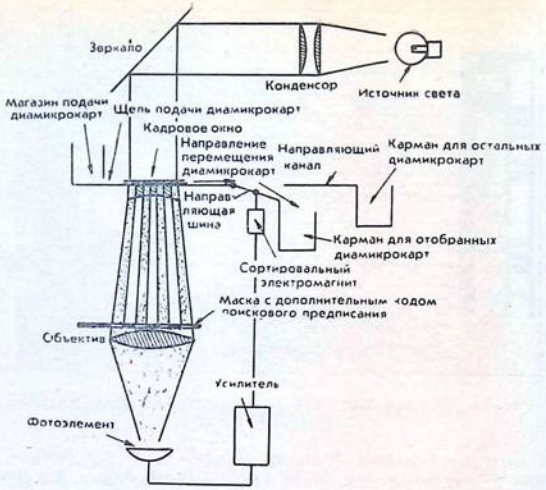



Рис. 68. Принцип действия селектора ИПС «Filmorex»

ческих ящиках размером 250 × 35 × 45 *мм*, а эти ящики — в специальном стеллаже. Емкость одного такого ящика 1,1 тыс. диамикрокарт, а емкость стеллажа 500 тыс. диамикрокарт.

Информационный поиск осуществляется при помощи селектора, принцип работы которого показан на рис. 68. Поисковое предписание записывается в дополнительном (комплементарном) коде, т. е. таким образом, что зачерненными оказываются те ячейки, которые прозрачны в основном коде дескриптора, и наоборот. При наложении дополнительного кода на основной строка кодового поля оказывается полностью маскированной. Отбор диамикрокарт может производиться одновременно лишь по трем дескрипторам.

Для этого изготовляется «негативная» маска кодов выбранных дескрипторов, представляющая собой карточку из зачерненной фотопленки, в которой при помощи клавишного перфоратора пробиты дополнительные коды этих дескрипторов. Такая маска вставляется в селектор и при сортировке диамикрокарт как бы накладывается на их кодовые поля. Если маска полностью экранирует кодовое поле микрокарты, то это означает, что в нем записаны коды дескрипторов, по которым производится информационный поиск. Совпадение кода с его «негативом» фиксируется тремя фотоэлементами, которые замыкают электрическую цепь сортировального магнита, направляющего диами-

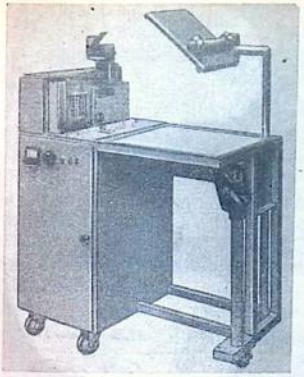


Рис. 69. Селектор ИПС «Filmorex»

Рис. 70. Увеличительно-печатающая установка ИПС «Filmorex»

крокарту в первый карман селектора. Диамикрокарты, не соответствующие поисковому предписанию, поступают во второй карман. Скорость работы селектора 600 *ДМК/мин*, емкость магазина подачи 900 ДМК.

В 1962 г. фирма «Filmorex» выпустила новую модель селектора (рис. 69), которая рассчитана на применение как диамикрокарт, так и рулонных 35-миллиметровых микрофильмов с изображениями диамикрокарт. Скорость работы этого селектора 400 *ДМК/мин*. Предусмотрена возможность замедления сортировки с тем, чтобы каждую диамикрокарту можно было визуально просматривать при помощи читального аппарата, встроенного в селектор. Скорость протяжки рулонного микрофильма в этом селекторе 0,5 *м/сек*, емкость бобины 30 или 90 *м*. Когда поисковый образ кадра (диамикрокарты) совпадает с введенным поисковым предписанием, микрофильм останавливается, и этот кадр проецируется на экран читального аппарата.

Функцию выходного устройства ИПС «Filmorex» выполняет увеличительно-печатающая установка (рис. 70), которая предназначена для автоматического изготовления увеличенных копий документов с диамикрокарт, а также с 35-миллиметровых рулонных и форматных микрофильмов. Копирование производится на рулонную фотобумагу шириной 110 *мм*. Запас бумаги 50 *мм*, кратность увеличения — постоянная (1 : 4). Установка выдает уже готовые к использованию копии документов. Ее производительность 40 диамикрокарт, или 3 *м/мин* рулонного микрофильма.

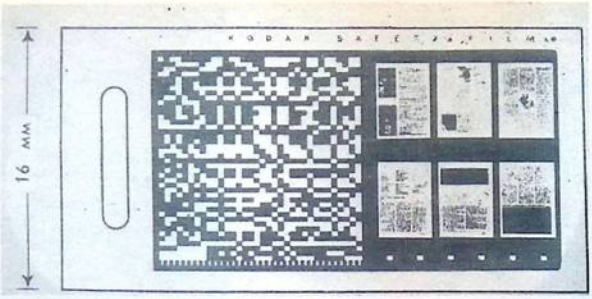



Рис. 71. Диамикрокарта типа «Minicard»

В настоящее время в практической эксплуатации находится около 20 комплектов ИПС «Filmorex», в том числе во Франции — 12, в Испании — 1, в Великобритании — 2 и в США — 1. Стоимость ИПС «Filmorex» 25 тыс. долл.

*Minicard*[30]. ИПС «Minicard» создана фирмой Eastman Kodak Co. (США) в 1953—1957 гг. по заказу Министерства ВВС США и практически используется в Пентагоне с весны 1957 г. На разработку этой ИПС было затрачено более 8 млн. долл. Из литературы известно, что в настоящее время в эксплуатации находятся четыре ИПС «Minicard», три из них используется военными ведомствами США (Министерством ВВС, Министерством обороны и Центральным разведывательным управлением), а четвертая — фирмой Eastman Kodak. Министерство ВВС США планировало,

---

[30] F. Bello. How to cope with information. «Fortune», 1960, v. 72, № 3, стр. 163—165, 167; A. W. Tyler, W. L. Myers, J. W. Kuipers. The application of the Kodak Minicard system to problems of documentation. «American Documentation», 1955, v. 6, № 1, стр. 18—30; J. W. Kuipers, A. W. Tyler, W. L. Myers. A Minicard system for documentary information. «American Documentation», 1957, v. 8, № 4, стр. 246—268; The Minicard system. A continuing development of Eastman Kodak Company. New York, Recordak Corp., 1958; A. M. Rahm. The Minicard system for high speed data retrieval. «MULL», 1960, June, стр. 63—67; W. L. Myers, G. L. Loomis. The Minicard film record as a common-language medium. В кн.: «Information retrieval and machine translation», pt. 1. New York, Interscience Publishers, 1960, стр. 575—624; T. C. Bagg, M. E. Stevens. Information selection system retrieving replica copies: A state-of-the-art report. Technical note 157. Washington, National Bureau of Standards, 1961, стр. 68—76; Current Research and Development in Scientific Documentation. N. 4, NSF-59-28, April, 1959, стр. 82—83; № 7, NSF-60-65, November, 1960.

в частности, ввести в ИПС «Minicard» более 20 млн. аэрофотоснимков.

В качестве носителей информации в ИПС «Minicard» используются диамикрокарты размером 16 × 32 *мм* (рис. 71). На такие диамикрокарты заносятся микрокопии документов вместе с их поисковыми образами. На одну диамикрокарту ИПС «Minicard» может быть занесено: а) до 12 кадров с микрокопиями документов форматом 22 × 36 *см* (кратность уменьшения 60 : 1); б) один кадр с микрокопией документа (карты, чертежи, схемы и т. п.) форматом 46 × 56 *см* (кратность уменьшения 38 : 1); в) один кадр с микрокопией документа форматом 46 × 33 *см* (кратность уменьшения 20 : 1); г) один кадр с микрокопией аэрофотоснимка размером 23 × 23 *см* или $^2/_3$ аэрофотоснимка размером 23 × × 46 *см* (кратность уменьшения 20 : 1).

Кодовое поле может занимать любую часть диамикрокарты и даже всю диамикрокарту. Если на диамикрокарту занесено 12 страниц документа размером 22 × 36 *см*, то емкость кодового поля минимальна и составляет 42 буквенно-цифровых знака (258 двоичных единиц); емкость кодового поля диамикрокарты, содержащей микрокопии 6 страниц такого же формата, составляет 259 буквенно-цифровых знаков (1591 двоичных единиц). Когда же для записи поискового образа документа используется вся диамикрокарта, то емкость кодового поля составляет 455 буквенно-цифровых знаков (2856 двоичных единиц). Схема записи информации в кодовом поле диамикрокарты показана на рис. 72. В одной колонке кодового поля содержится 43 ячейки и записывается 7 буквенно-цифровых знаков (по 6 двоичных единиц на знак). Последняя, 43-я ячейка каждой колонки используется для контроля по четности. Минимальное число колонок в кодовом поле[31] диамикрокарты — 6, максимальное — 68. В ИПС «Minicard» возможно применение любого числа диамикрокарт-продолжений.

На рис. 72 видно, что в правой части диамикрокарты имеется поперечный щелевой вырез. Этот вырез предназначен для нанизывания диамикрокарт на специальные плоские штыри, которые применяются для переноса диамикрокарт из одного функционального блока ИПС «Minicard» в другой. Длина штыря 406 *мм*, емкость 2 тыс. диамикрокарт.

В комплект оборудования ИПС «Minicard» входят следующие установки и устройства:

1. *Микрофильмирующая установка.* Она предназначена для съемки на 16-миллиметровый микрофильм документов и их поисковых образов. Поисковые образы документов вводятся в микрофильмирующую установку при помощи перфоленты. Установка

[31] При репродуцировании диамикрокарт к их кодовым полям может быть добавлена седьмая колонка.

12. *Быстродействующий ленточный перфоратор* (скорость работы 6 тыс. знаков в 1 мин.) — или флексорайтер (скорость работы 600 знаков в 1 мин).

13. *Шкафы для хранения диамикрокарт.* Диамикрокарты хранятся в алюминиевых блоках (по три блока в выдвижном ящике шкафа). Каждый такой блок имеет 50 магазинов [32]; емкость одного магазина 2 тыс. диамикрокарт. Общая емкость шкафа с тремя выдвижными ящиками 900 тыс. диамикрокарт, на которых может содержаться до 10,8 млн. кадров с микрокопиями документов форматом $22 \times 36$ *см*. Размер шкафа $380 \times 760 \times 1270$ *мм*. Следует указать, что для хранения такого массива документов потребовалось бы более 500 стандартных шкафов и рабочая площадь 230 $м^2$.

Ввод документов в ИПС «Minicard» производится следующим образом. Сначала документы проходят соответствующую регистрацию, и на них проставляются регистрационные номера.

Предметизатор, производящий индексирование документа, переносит регистрационный номер этого документа в специальный формуляр, в котором он записывает также поисковый обзор данного документа. Заполненный формуляр передается оператору, который при помощи флексорайтера пробивает содержимое этого формуляра на перфоленте. Поскольку флексорайтер одновременно воспроизводит кодируемый текст в печатном виде, то оператор визуально контролирует правильность записи. Далее перфолента с номерами документов и их поисковыми образами подается в вводное устройство микрофильмирующей установки. Туда же поступают оригиналы вводимых в ИПС «Minicard» документов. Оператор микрофильмирующей установки сравнивает регистрационный номер поступившего на съемку документа с номером того документа, поисковый образ которого введен в установку. Выше уже указывалось, что в микрофильмирующей установке типа CD3A этот номер в дешифрованном виде высвечивается на специальном контрольном экране. При совпадении этих двух номеров производится фотосъемка документа, во время которой на 16-миллиметровый микрофильм автоматически заносится и код поискового образа документа.

После фотосъемки кассета с экспонированной пленкой подается в проявочную машину, которая производит полную обработку (проявление, фиксирование, промывку и сушку) бобины микрофильма (61 *м*) примерно за 40 мин. Готовый негативный микрофильм проходит технический контроль и подается в гильотину, которая автоматически разрезает его на диамикрокарты. Эти негативные диамикрокарты (создатели ИПС «Minicard» называют их негативными диамикрокартами «первого поколения» и обозначают такие диамикрокарты символом 1N) поступают

[32] Применяются также блоки, имеющие по 10 магазинов.

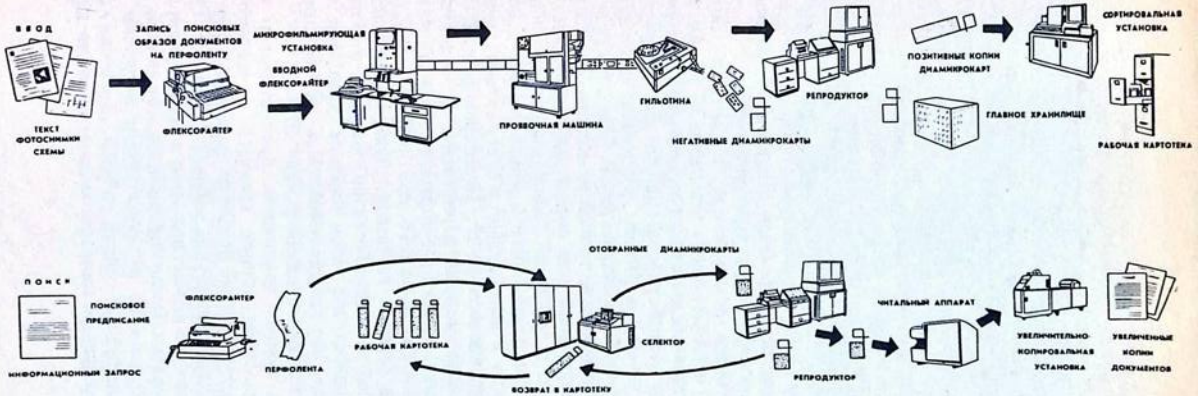



Рис. 73. Принцип действия и основные устройства ИПС «Minicard»

в приемный магазин гильотины, где они нанизываются на транспортировочные штыри. Заполненные штыри с диамикрокартами 1N поступают в главное хранилище (рис. 73).

Необходимо отметить, что диамикрокарты 1N для информационного поиска не используются и хранятся в ИПС «Minicard» как оригиналы. Рабочие картотеки ИПС комплектуются из так называемых позитивных диамикрокарт «второго поколения» (они обозначаются символом 2P), изготовляемых путем контактного копирования диамикрокарт 1N перед их передачей в главное хранилище. При копировании диамикрокарт 1N еще раз производится контроль соответствия между поисковым образом и номером документа, которому присвоен поисковый образ. Для этого репродуктор выдает в дешифрованном печатном виде и в виде перфоленты поисковый образ, записанный на диамикрокарте, и номер документа, занесенного на эту диамикрокарту. Полученную перфоленту можно автоматически сравнить с исходной перфолентой, которая использовалась для ввода поисковых образов документов в микрофильмирующую установку.

При копировании диамикрокарт 1N ЭЦВМ репродуктора автоматически определяет число экземпляров диамикрокарт 2P, которое необходимо для включения данной диамикрокарты во все соответствующие секции рабочей картотеки. Индексы (шифры) таких секций рабочей картотеки заносятся в фиксированную часть кодовых полей диамикрокарт 2P. Это делается для того, чтобы при информационном поиске не нужно было производить сортировку диамикрокарт всех секций рабочей картотеки.

Экспонированный рулонный микрофильм из репродуктора поступает на проявление; после проявления он проходит технический контроль и далее подается в гильотину, разрезающую микрофильм на диамикрокарты 2P. Диамикрокарты 2P при помощи транспортировочных штырей переносятся в селектор-сортировальную установку, которая раскладывает их по индексам секций и подсекций рабочей картотеки. После сортировки диамикрокарты 2P поступают в рабочую картотеку. Диамикрокарты 1N после репродуцирования раскладываются по номерам документов (или датам их поступления) и в таком порядке хранятся в главном хранилище. Далее они используются лишь для воспроизведения приходящих в негодность диамикрокарт 2P.

Для поиска нужных документов предметизатор составляет поисковое предписание, которое вводится в селектор-сортировальную установку. Туда же подаются диамикрокарты 2P из соответствующей секции или секций рабочей картотеки. Отобранные селектором диамикрокарты 2P подаются в репродуктор, копирующий их на микрофильм. После информационного поиска диамикрокарты 2P сортируются по их картотечным индексам и возвращаются в соответствующие секции рабочей картотеки, а экспонированный в репродукторе микрофильм поступает на проявление.

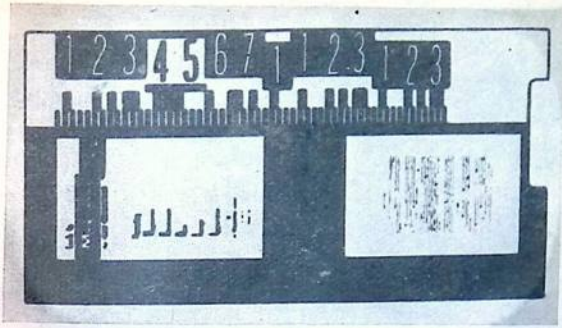


Рис. 74. Диамикрокарта ИПС «MEDIA»

Проявленный негативный микрофильм проходит технический контроль и разрезается гильотиной на диамикрокарты 3N. В зависимости от условий заказа диамикрокарты 3N направляются заказчику, монтируются в апертурные карточки или передаются на увеличение и печать. Перед увеличением диамикрокарты могут быть просмотрены заказчиком при помощи консольного читального аппарата. Среднее время поиска нужных документов в ИПС «Minicard» (включая время изготовления диамикрокарт 3N) обычно составляет около 10 мин.

Стоимость ИПС «Minicard» — около 2 млн. долл.

*MEDIA* [33] — это электронная система для поиска микрокопий документов по известным адресам их хранения [34]. Система MEDIA создана фирмой Magnovox Co. (США) в 1960 г. Носитель информации в системе MEDIA — позитивная диамикрокарта размером 16×32 *мм* (рис. 74). Емкость диамикрокарты — две страницы документа формата 23×38 *см* или три страницы формата 22×28 *см*. Кратность уменьшения — 30 : 1. Адресный шифр документа в виде десятичных цифр и соответствующих им двоич-

---

[33] MEDIA — Magnavox Electronic Data Image Apparatus — «Электронный аппарат фирмы Magnavox Co. для поиска графической информации».

[34] Film cards are key to data retrieval. «Business Automation», 1962, v. 8, № 5, стр. 46; Faster ways to store, retrieve data. «Plat Administration and Engineering», 1962, v. 22, № 12, стр. 32—33; Graphic information storage and retrieval system. «Data Processing» (Engl.), 1963, v. 5, № 5, стр. 322—324; D. D. Jenkins. Magnetic indexing, microfilm storage and information retrieval. «Proceedings of the 11th Annual Meeting and Convention of the National Microfilm Association», Annapolis, National Microfilm Association, 1962, стр. 205—212.

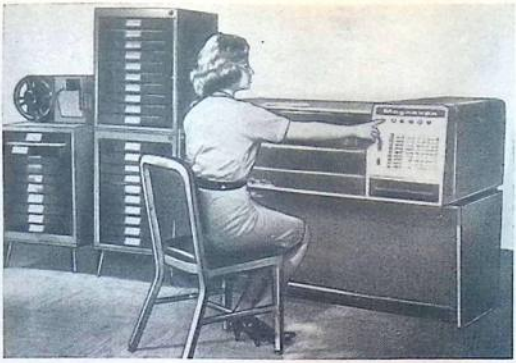


Рис. 75. Селектор-репродуктор ИПС «MEDIA»

ных кодов записывается вдоль верхнего края диамикрокарты. Максимальная емкость кодового поля — 17 десятичных цифр (69 двоичных знаков). Цифровой шифр на диамикрокарте легко читается невооруженным глазом.

Система MEDIA состоит из трех устройств: установки для микрофильмирования (с кодирующим устройством), гильотины для разрезки микрофильма на диамикрокарты и селектора-репродуктора. Установка для микрофильмирования — вертикального типа. Ввод адресного шифра микрофильмируемого документа производится при помощи клавиатуры. Адресный шифр высвечивается на кодовой линейке, которая фотографируется вместе с документом. Микрофильмирование производится на 16-миллиметровую неперфорированную пленку. Емкость бобины — 30,5 *м*. После проявления микрофильм контратипируется; полученная копия при помощи гильотины разрезается на диамикрокарты. Скорость работы гильотины — 4 *ДМК/сек*.

Диамикрокарты хранятся в пронумерованных обоймах, а обоймы — в пронумерованных выдвижных ящиках специальных шкафов (по 100 обойм в ящике). Емкость одной обоймы — до 200 диамикрокарт. В обойме диамикрокарты располагаются в произвольном порядке. В одном шкафу размером 635 × 490 × 510 *мм* может храниться до тысячи таких обойм, т. е. 200 тыс. диамикрокарт.

Если необходима, например, диамикрокарта с документом № 1234567, то оператор находит в шкафу обойму № 12345 и вставляет ее в селектор-репродуктор (рис. 75). Затем на клавиатуре селектора-репродуктора он набирает две последние цифры но-

мера — 67 и включает установку. В селекторе-репродукторе диамикрокарты автоматически извлекаются из обоймы (рис. 76) и сортируются со скоростью 10 *ДМК/сек*. С диамикрокарты, содержащей искомый документ, в случае необходимости может быть немедленно изготовлена увеличенная до нормального размера копия. Возможен и визуальный просмотр диамикрокарты при помощи читального аппарата. После сортировки диамикрокарты автоматически укладываются в исходную обойму. Продолжительность поиска диамикрокарты в обойме (вместе с изготовлением увеличенной копии найденного документа) — менее одной минуты. Стоимость ИПС MEDIA — 43 тыс. долл.

*Magnacard* [35]. Система типа «Magnacard» создана американской фирмой Magnavox Co. в 1957 г. по заказу Министерства ВВС США. В качестве носителя информации в этой ИПС используется магнитная карта размером 25,4×76,2 *мм*. Магнитные карты изготовляются из майларовой пленки толщиной 0,127 *мм*, на которую с одной стороны нанесен ферромагнитный слой (окись железа) толщиной 0,013 *мм*. Для предохранения от повреждений на этот слой нанесена майларовая пленка толщиной 0,013 *мм*. Специалисты утверждают, что по магнитным свойствам такие карты не уступают лучшим магнитным лентам. Запись информации на магнитных картах производится при помощи головок, которые не контактируют с поверхностью карт. Благодаря этому магнитная карта допускает не менее 20 тыс. прогонов через систему. Информационная емкость магнитной карты — 756 буквенных или 1134 цифровых знаков (5060 двоичных единиц). Запись и считывание информации с магнитной карты производится со скоростью 90 тыс. знаков/сек.

Рис. 76. Обойма с диамикрокартами ИПС «MEDIA»

На рис. 77 приведена схема записи информации на магнитной карте. Из этого рисунка видно, что магнитная карта имеет

[35] R. L. Laurent. Information storage and retrieval. «Missile Design and Development», 1960, № 10, стр. 56—58; R. M. Hayes. The Magnacard system. В кн.: «Information retrieval and machine translation», pt. 1. New York, Interscience Publishers, 1960, стр. 563—574; R. E. Westgard. The Magnacard system. «Datamation», 1961, v. 7, № 7, стр. 42—43; R. L. Laurent. Magnacard. A unique approach to the problem of combining massive data storage with rapid on line access. «Data Processing», 1962, v. 4, № 4, стр. 33—38; Magnacard. «Data Processing», (Engl.), 1962, v. 4, № 1, стр. 16—25.

Рис. 77. Схема записи информации на магнитной карте типа «Magnacard»

20 дорожек, в которых поперек карты (колонками) записывается буквенная и цифровая информация. Две нижние дорожки используются для вспомогательных целей: одна — для контроля по четности, а вторая — для синхронизирующих импульсов. В одной колонке можно записать три буквенно-цифровых знака (по 6 двоичных разрядов на знак). На одной магнитной карте умещается 253 таких колонки (плотность записи — 4 двоичных единицы на 1 *мм*). Крайняя слева колонка используется для записи контрольной суммы. Благодаря такой системе двойного контроля обеспечивается высокая точность считывания информации с магнитных карт.

Видеомагнитная карта типа «Magnavue» (рис. 78) имеет такие же размеры, как и магнитная карта описанного выше типа,

Рис. 78. Схема записи информации на видеомагнитной карте типа «Magnavue»

по отличается от нее по своей конструкции. Видеомагнитная карта состоит из трех слоев майларовой пленки. Первый слой, имеющий толщину 0,025 *мм*, предохраняет от повреждений диамикрокарту с позитивным микроизображением. Толщина второго слоя — 0,1 *мм*; на него нанесено ферромагнитное покрытие (окись железа) толщиной 0,013 *мм*. В этой пленке сделан калиброванный вырез для диамикрокарты. Сверху ферромагнитное покрытие и диамикрокарта защищены третьим слоем майларовой

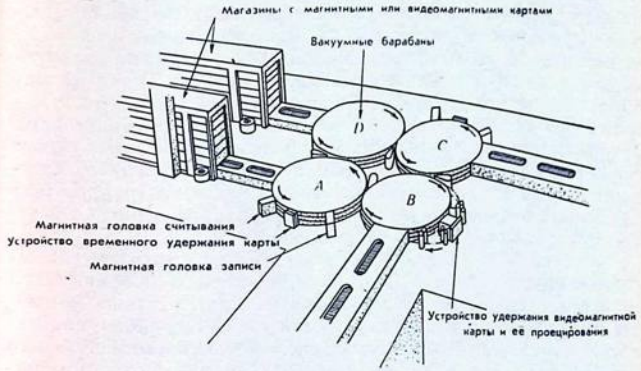



Рис. 79. Схема верхней панели сортировальной установки ИПС «Magnacard»

пленки толщиной 0,013 *мм*. Ферромагнитное покрытие занимает $^2/_3$ длины видеомагнитной карты. Остальная часть видеомагнитной карты занята микроизображением документа. Емкость кодового поля видеомагнитной карты — 450 буквенных или 675 цифровых знаков (3 тыс. двоичных единиц). Запись информации в этом кодовом поле производится по той же системе, что и на магнитных картах типа «Magnacard». Видеомагнитная карта описанной конструкции допускает не менее 200 тыс. прогонов через систему [36].

Центральным блоком системы «Magnacard» является сортировальная установка с четырьмя вакуумными барабанами. Схема верхней панели такой установки приведена на рис. 79. Магнитные или видеомагнитные карты поступают в сортировальную установку из горизонтальных магазинов. Емкость каждого такого магазина — до 3 тыс. карт. Сортировка магнитных и видео-

[36] R. L. Laurent. Combined magnetic and graphic store. В кн.: «Large-Capacity Memory Techniques for Computing Systems». New York—London, Macmillan, 1962, стр. 137—147.

магнитных карт производится при помощи системы вакуумных барабанов. Вакуумный барабан представляет собой полый цилиндр диаметром 203 *мм* и высотой 25,4 *мм*, вращающийся вокруг вертикальной оси со скоростью 12 *оборотов/сек*. На боковой поверхности барабана имеются две кольцевые щели. Из внутренней полости барабана постоянно откачивается воздух. Поэтому у поверхности барабана создается перепад давлений. Воздух, попадающий внутрь барабана через его боковые щели, затягивает гибкие магнитные или видеомагнитные карты и надежно удерживает их на боковой поверхности барабана. Скорость подачи карт из магазина — 90 *шт/сек*, линейная скорость перемещения вращающимся барабаном — 7,6 *м/сек*. Вакуумные барабаны почти контактируют один с другим. В тех местах, где расстояние между барабанами является минимальным, установлены по два пневматических клапана, при помощи которых осуществляется перенос магнитных и видеомагнитных карт с одного барабана на другой. Перенос карты на другой барабан осуществляется в нужный момент времени струи воздуха подачи в нужную кромку карты. Каждый клапан обеспечивает перемещение только в одном направлении.

Барабаны *A* и *C* снабжены магазинами. Барабаны *A*, *C* и *D* (рис. 79) имеют по одному магазину, а барабан *B* — два магазина. Барабаны *A* и *C* снабжены устройствами записи и считывания, а также устройствами задержки карт. Если такое устройство снимает с поверхности вакуумного барабана *A* магнитную карту и видеомагнитную карту с поверхности барабана *C*, то карта на этот барабан автоматически возвращается на поверхность барабана лишь после того, как карта возобновляется устройством для задержки видеомагнитных карт. Такие устройства позволяют использовать их изображений на большой экран. В случае необходимости с видеомагнитной связи; в случае увеличенные копии изображений. Видеомагнитными картами хранятся в специальных хранилищах (рис. 80). В системе применяются три типа таких хранилищ: вертикальные, горизонтальные и горизонтальные. В вертикальных хранилищах карты располагаются в специальную клетку друг над другом. Емкость одного такого хранилища — 50 магазинов производится путем установки. Когда нужный магазинами относительно вакуумных барабанов, подводится к поверхности одного вертикального хранилища с 22 млн. буквенно-цифровых установке может быть одновре-

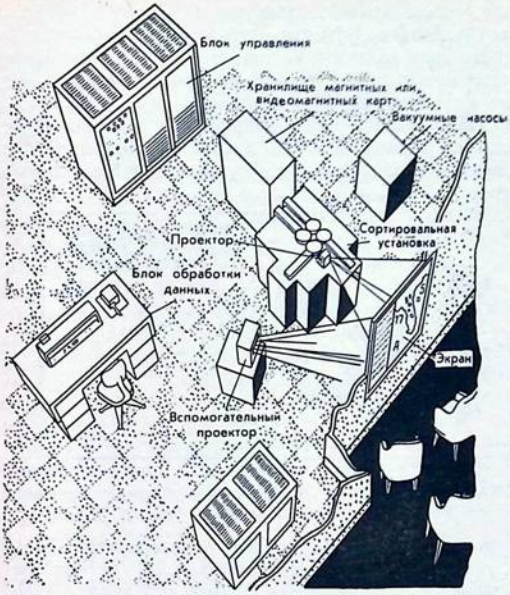



Рис. 80. Основные блоки и устройства ИПС «Magnacard»

менно присоединено пять таких хранилищ. Среднее время смены магазина — 3 сек., максимальное время — 6 сек.

К сортировальной установке можно подключить также вспомогательное хранилище, имеющее 10 рядов по 10 карманов в каждом ряду. Емкость каждого кармана — 200 карт. Выбор нужного кармана производится путем соответствующего перемещения всего блока хранилища по вертикали и горизонтали. Из кармана хранилища магнитные и видеомагнитные карты перемещаются в обычный магазин и из него подаются на сортировку. Емкость вспомогательного хранилища — 20 тыс. карт, время подачи карт на сортировку из любого кармана — 5 сек.

Для хранения очень больших массивов информации (до 450 тыс. карт) предназначены горизонтальные хранилища. Горизонтальное хранилище составлено из 15 вертикальных хранилищ, каждое из которых содержит 10 магазинов. Выбор нужного магазина производится путем горизонтального перемещения вертикальных хранилищ и подъема (опускания) клетки вертикального хранилища, в котором находится искомый магазин.

К одной сортировальной установке могут быть присоединены два таких хранилища. Время доступа в такое хранилище — около 30 сек. Имеются сведения о том, что созданы горизонтальные хранилища емкостью 600 тыс. и 900 тыс. карт.

Работой сортировальной установки может управлять универсальная ЭЦВМ любого типа. В прототипе системы «Magnacard» для этой цели применялась ЭЦВМ CDC 160 фирмы Control Data Corp. Ввод информации в систему производится при помощи

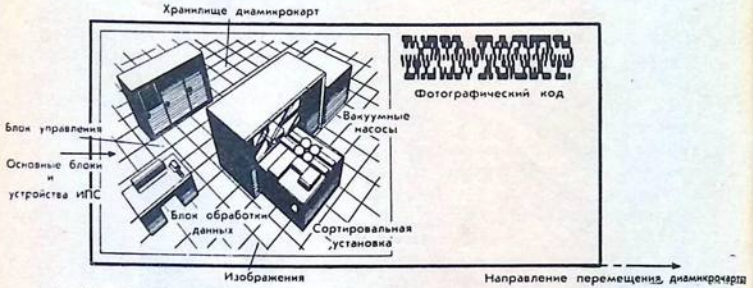



Рис. 81. Диамикрокарта ИПС «DARE» с изображением основных блоков и устройств этой системы

перфоленты и перфокарт. В комплект описываемого оборудования входят также устройства «Magnascriber» и «Veriscriber», предназначенные соответственно для статической записи информации непосредственно на магнитных и видеомагнитных картах и для верификации этой записи. Устройства «Magnascriber» и «Veriscriber» имеют такую же клавиатуру, как и клавишные перфораторы, и не превышают размеров пишущей машины. При помощи таких устройств на магнитной или видеомагнитной карте можно записать до 128 знаков.

В настоящее время фирма Magnovox Co. на базе описанного выше оборудования разработала специализированную ИПС «DARE», предназначенную для хранения и поиска технической документации[37]. Носителями информации в этой ИПС являются диамикрокарты размером $35 \times 76$ *мм*. Диамикрокарты изготовляются из диазопленки с майларовой подложкой. Схема такой диамикрокарты приведена на рис. 81. В левой части диамикрокарты помещается микрокопия технического документа, а в правой части записывается (в виде фотографического кода) поисковый образ документа. Емкость кодового поля диамикрокарты — 80 буквенно-цифровых знаков (480 двоичных единиц).

[37] R. L. Laurent. Magnavue system for automated engineering drawing storage and retrieval. «Automation and Scientific Communication». Washington, American Documentation Institute, 1963, стр. 91—92.

Диамикрокарты хранятся в магазинах емкостью 3 тыс. карт каждый. ИПС снабжена хранилищем с быстрой выборкой; емкость хранилища — 225 магазинов. Когда производится поиск диамикрокарты в одном магазине, механизм выборки может устанавливать в исходное положение следующий магазин. Считка поисковых образов производится фотооптической головкой, встроенной в сортировальную установку. Скорость сортировки диамикрокарт — 90 *шт/сек*. Один из вакуумных барабанов снабжен устройством для задержки и диазокопирования диамикрокарт. Экспонированная диазопленка автоматически монтируется в апертурные перфокарты и затем проявляется. С диамикрокарты может быть изготовлена одна или несколько копий. Кроме того, это устройство автоматически переносит с диамикрокарты на апертурную перфокарту поисковый образ документа.

Изготовление диамикрокарт производится при помощи специальной установки, в которую вводятся апертурные перфокарты с микрокопиями технических документов (главным образом — чертежей) и перфокарты с поисковыми образами этих документов. Рекопирование микроизображений производится на 35-миллиметровую рулонную диазопленку контактным способом. После проявления эта пленка разрезается на диамикрокарты, которые затем укладываются в магазины.

Обычно для выполнения заказов, поступивших за день, производится просмотр всего хранилища, на что уходит около 6 часов. Когда поступают срочные заказы, нормальная процедура нарушается и требуемая копия выдается не более чем через минуту после ввода в ИПС адресного шифра необходимого документа.

*Walnut* [38]. ИПС «Walnut», что в переводе с английского означает «грецкий орех», создана американской фирмой International Business Machines Corp. по заказу Центрального разведывательного управления США (ЦРУ). Первые сведения об этой ИПС попали на страницы американской печати в июле 1961 г., хотя ее разработка велась с 1958 г. В этой связи следует отметить следующее весьма многозначительное обстоятельство. По-видимому, с начала 1959 г. в ЦРУ находится в эксплуатации одно-

---

[38] Large-capacity document storage and retrieval system. «Engineer», 1961, v. 212, № 5508, стр. 291—292; CIA's film memory to store 990 000 pages. «Control Engineering», 1961, v. 8, № 7, стр. 23; A Walnut for information retrieval. «Data Processing», 1961, v. 3, № 8, стр. 26—27; P. D. Bradshaw. The Walnut system: a large capacity document storage and retrieval system. «American Documentation», 1962, v. 13, № 3, стр. 270—275; W. E. Bushop. Information storage/retrieval. «Electronics», 1962, v. 35, № 26, стр. 54—55; R. W. Porter. A large-capacity document storage and retrieval. В кн.: «Large-capacity memory techniques for computing systems». New York, Macmillan, 1962, стр. 351—360; N. A. Vogel. Walnut document storage and retrieval system. «Proceedings of the 11th Annual Meeting and Convention of the National Microfilm Association», Annapolis, National Microfilm Association, 1963, стр. 27—39.

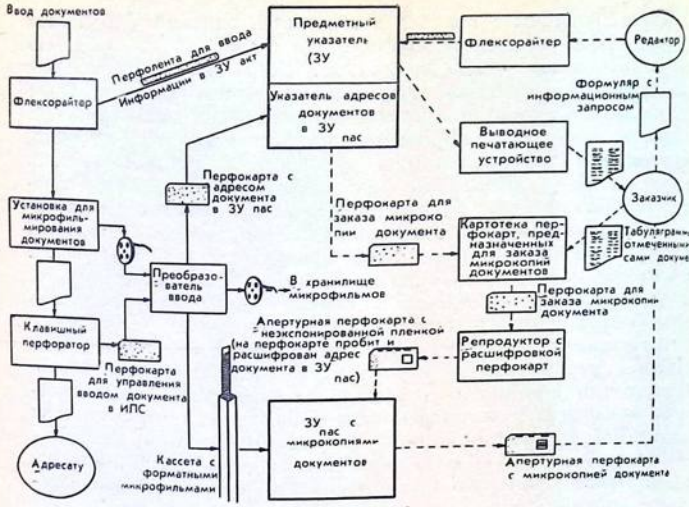



Рис. 82. Структурная схема ИПС «Walnut»

контурная ИПС «Minicard». Тот факт, что ЦРУ заказало фирме International Business Machines Corp. разработку новой ИПС, уже имея в своем распоряжении ИПС «Minicard», не может свидетельствовать в пользу высоких эксплуатационных качеств ИПС «Minicard».

В ИПС «Walnut» применена схема с двумя ЗУ — активным и пассивным. В качестве $ЗУ_{акт.}$ в ней используется ЭЦВМ IBM-1440 с несколькими внешними накопителями типа IBM-1301, выполненными на магнитных дисках. В этих накопителях записываются поисковые образы документов, микрокопии которых хранятся в $ЗУ_{пас.}$ ИПС. Доступ в накопители — произвольный, время обращения — не более 2 сек. Индексирование документов, вводимых в ИПС «Walnut», производится квалифицированными специалистами. Поисковые образы документов вводятся в $ЗУ_{акт.}$ на перфоленте. Структурная схема ИПС «Walnut» показана на рис. 82.

$ЗУ_{пас.}$ IBM-9603 имеет модульную конструкцию. Не превышая по размерам письменный стол, каждый модуль $ЗУ_{пас.}$ вмещает 990 тыс. страниц документов, что по объему эквивалентно около 3 тыс. книг средней величины. Всего в ИПС «Walnut» может быть применено до 100 и более таких модулей. Доступ в $ЗУ_{пас.}$ — произвольный, максимальное время обращения

(вместе со временем, затрачиваемым на рекопирование требуемого документа) не превышает 5 сек., среднее время обращения — 1,5 сек. Носителем информации в $ЗУ_{нас.}$ является форматный микрофильм размером 23 × 394 *мм*, на котором в три ряда размещаются 99 кадров размером 6×10 *мм* каждый с микрокопиями документов. Каждый кадр отводится под микрокопию одной страницы документа форматом 20×36 *см*. Применяемая кратность уменьшения — 35 : 1. Микрокопии изготовляются сухим способом на фотопленке типа «кальфакс», разрешающая способность которой составляет 550 линий на 1 *мм*. В фотопленке типа «кальфакс» светочувствительным слоем является полимерный носитель, в котором диспергировано соответствующее диазосоединение. Под действием УФ-лучей молекулы этого диазосоединения разлагаются и внутри носителя (за счет давления газообразных продуктов реакции) образуются локальные области внутренних напряжений. Проявление скрытого изображения производится путем нагревания экспонированной фотопленки до 120° С, в результате которого локальные области внутренних напряжений превращаются в области упорядоченной («кристаллической») структуры, отличающиеся по показателям преломления от остальных областей, где не было фоторазложения. Возникшее таким образом светорассеивающее (а не светопоглощающее, как в галоидносеребряных слоях) изображение доступно непосредственному визуальному наблюдению.

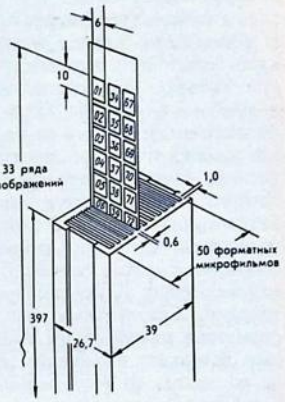


Рис. 83. Кассета с форматными микрофильмами ИПС «Walnut»

Форматные микрофильмы хранятся в пластмассовых кассетах по 50 шт. в кассете, а эти кассеты — в радиальных магазинах вертикального поворотного барабана модуля $ЗУ_{нас.}$. Всего в барабане модуля имеется 40 таких магазинов емкостью по 5 кассет каждый. Ввод в $ЗУ_{нас.}$ адресов требуемых документов осуществляется при помощи апертурных перфокарт, в которые вклеен прямоугольник неэкспонированной пленки типа «кальфакс». На эту пленку с небольшим увеличением (до кратности 27,2 : 1) рекопируется по четыре кадра отбираемых документов. Рекопирование производится полностью автоматически в УФ-лучах; проявление экспонированной пленки осуществляется путем ее теплового облучения. Продолжительность проявления —

около 0,5 сек. Таким образом, основные форматные микрофильмы с микрокопиями документов никогда не покидают $ЗУ_{нас.}$. Схема кассеты с форматными микрофильмами показана на рис. 83.

Адрес документа в $ЗУ_{нас.}$ ИПС «Walnut» состоит из номера модуля $ЗУ_{нас.}$, номера кассеты в этом модуле, порядкового номера форматного микрофильма в кассете, номера кадра с микрокопией первой страницы данного документа на этом форматном микрофильме и числа страниц в документе. После считывания адреса с апертурной перфокарты барабан $ЗУ_{нас.}$ путем поворота и радиального сдвига в кратчайшем направлении устанавливает нужную кассету в положение выборки. Затем специальное устройство юстирует положение кассеты в блоке выборки, зажимной механизм этого блока захватывает нужный форматный микрофильм и выдвигает его из кассеты настолько, чтобы первый кадр микрокопии требуемого документа оказался в положении рекопирования. Одновременно производится сдвиг оптической системы, которая устанавливается напротив одной из трех колонок форматного микрофильма. В блоке выборки форматные микрофильмы перемещаются по направляющим, которые по размерам точно совпадают с пазами кассеты. Смещение пазов кассеты относительно направляющих блока выборки предотвращается системой блокировки. Специальным фотоэлементом контролируется, чтобы форматные микрофильмы после рекопирования заняли в кассете свое начальное положение. Если микрофильм вошел в кассету не до конца, то этот фотоэлемент препятствует разблокировке барабана. Схема работы модуля $ЗУ_{нас.}$ показана на рис. 84.

Третьим основным функциональным узлом ИПС «Walnut» является преобразователь ввода IBM-9403. Для ввода в ИПС

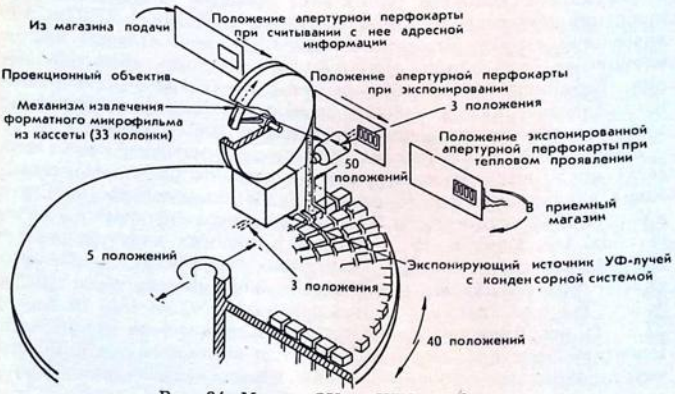



Рис. 84. Модуль $ЗУ_{нас.}$ ИПС «Walnut»

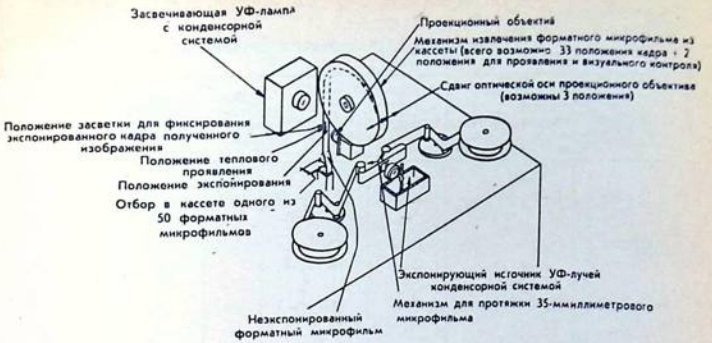



Рис. 85. Верхняя панель преобразователя IBM-9403 ИПС «Walnut»

«Walnut» документы сначала микрофильмируются на 35-миллиметровую перфорированную галоидносеребряную пленку. Кратность уменьшения при микрофильмировании — 15,8 : 1. После микрофильмирования на каждый документ заводится одна перфокарта, на которой пробивается порядковый номер документа, присвоенный ему при индексировании, а также число страниц в этом документе. В преобразователь ввода подаются бобина микрофильма, массив перфокарт, подобранных по порядковым номерам документов на этом микрофильме, а также кассета с неэкспонированной форматной пленкой типа «кальфакс». Преобразователь ввода автоматически переносит микрокопии документов на форматную фотопленку (с уменьшением 2,2 : 1), проявляет экспонированные кадры, вырабатывает адреса документов в $ЗУ_{нас.}$ и заносит эти адреса на форматные микрофильмы и перфокарты. Перфокарты далее используются для ввода в $ЗУ_{акт.}$ адресов документов в $ЗУ_{пас}$ и информации о числе страниц в каждом документе. Максимальная производительность преобразователя ввода — 1500 *кадров/час*. Схема верхней панели преобразователя IBM-9403 показана на рис. 85.

Информационный поиск в системе «Walnut» производится следующим образом. Заказчик заполняет специальный формуляр, в котором указывает возможно больше ключевых слов, характеризующих информационное содержание нужных ему документов. Заполненный формуляр редактируется специалистом; после этого четко сформулированное поисковое предписание пробивается на перфоленте для ввода в $ЗУ_{акт.}$. $ЗУ_{акт.}$ выдает табуляграмму, на которой приводятся адреса всех документов, содержащих требуемую информацию, и полные наборы дескрипторов, описы-

| Название | Тип поиска | Страна | Год создания | Тип носителя и его размеры, мм | Емкость в кадрах | Кратность уменьшения документов | Емкость кодового поля в двоичных единицах |
|---|---|---|---|---|---|---|---|
| Filmorex | С | Франция | 1950 | ДМК *, Н 35 × 60 | 1 | 4 : 4 + + 30 : 1 | Не более 400 |
| Minicard | С | США | 1957 | ДМК, П 16 × 32 | 0+12 | 60 : 1; 38 : 1; 20 : 1 | 258 ÷ 2856 |
| Magnacard | С, А | США | 1957 | МК, — 25,4 × 76,2 | — | — | 5060 |
| Magnavue | С, А | США | 1957 | ВМК, П 25,4 × 76,2 | 1 | — | 3020 |
| DARE | С, А | США | 1963 | ДМК, П 35,0 × 76,2 | 1 | — | 480 |
| MEDIA | А | США | 1960 | ДМК, П 16 × 32 | 2; 3 | 30 : 1 | 69 |
| Walnut | С, А | США | 1961 | ФМФ, П 23 × 394 | 99 | 35 : 1 | — |

* В последней модели селектора «*Filmorex*» возможно также применение рулонного ДМК — диамикрокарта; МК — магнитная карта; Н — негативная; П — позитивная; С —

вающих информационное содержание этих документов. Одновременно выводное устройство $ЗУ_{акт.}$ заполняет на каждый отобранный документ перфокарту, пробивая на ней адрес этого документа в $ЗУ_{пас.}$, число страниц в документе и его регистрационный номер. Заказчик просматривает табуляграмму и помечает на ней те документы, которые, по его мнению, в наибольшей степени отвечают его запросу и с которыми он хотел бы ознакомиться. После этого адресные перфокарты всех отмеченных заказчиком документов подаются в репродуктор IBM-528, пробивающий на апертурных перфокартах адреса этих документов в $ЗУ_{пас}$, число страниц в каждом документе и шифр заказчика. Емкость одной апертурной перфокарты — четыре кадра микрокопии. Поэтому если требуемый документ содержит более четырех страниц, то репродуктор автоматически изготовляет необходимое число апертурных перфокарт-продолжений. Эти апертурные перфокарты поступают в соответствующий модуль $ЗУ_{пас.}$, где на них автоматически рекопируются требуемые документы. В этой связи следует отметить, что ИПС «Walnut» можно легко модифицировать таким образом, чтобы информация из $ЗУ_{акт.}$ поступала непосредственно в $ЗУ_{пас.}$ без промежуточного корректирования.

**устройств с дискретными носителями информации**

| Скорость сортировки, *карт/сек* | Среднее время поиска | Емкость одного хранилища, число носителей | Способ вывода найденных документов | | | | Ориентировочная стоимость, тыс. долл. |
|---|---|---|---|---|---|---|---|
| | | | на экран | контактные копии | апертурные карты | увеличенные копии на бумаге | |
| 4; 6 | 5 мин. | $5 \times 10^5$ | + | — | — | + | 25 |
| 20 | 10 мин. | $9 \times 10^5$ | — | + | + | + | 750—2000 |
| 90 | 30 сек. | $2 \times 10^4 \div 1{,}8 \times 10^6$ | — | — | — | — | 250 |
| 90 | 30 сек. | $2 \times 10^4 \div 1{,}8 \times 10^6$ | + | + | — | + | 400 |
| 90 | 1 мин. | $6{,}75 \times 10^5$ | — | — | + | — | |
| 10 | 1 мин. | $20^5$ | — | — | — | + | 43 |
| — | 30 сек. | $9{,}9 \times 10^5$ | — | — | + | — | 500—1000 |

35-миллиметрового микрофильма. А — поиск по адресу; ПК — видеомагнитная карта; поиск по содержанию; ФМФ — форматный микрофильм.

*Verac-903*. Система «Verac-903» представляет собой $ЗУ_{нас.}$ матричного типа [39], в котором может храниться до 1 млн. микрокопий документов формата не более 24 × 36 *см*. Это $ЗУ_{нас.}$ разработано фирмой Avco Corp. (США) в результате научно-исследовательских и опытно-конструкторских работ, которые проводятся ею с конца 1956 г. до настоящего времени. $ЗУ_{нас.}$ «Verac-903» предназначено для двухконтурных ИПС, в которых может быть применено $ЗУ_{акт.}$ любого типа. $ЗУ_{нас.}$ «Verac-903» состоит из четырех подсистем: а) установки для микрофотографирования документов; б) собственно $ЗУ_{нас.}$; в) системы вывода и г) электрографической установки для изготовления увеличенных копий документов.

Носителем информации $ЗУ_{нас.}$ «Verac-903» является матричная плата размером 198 × 185 *мм*, покрытая фотоэмульсией с разрешающей способностью не менее 1500 линий на 1 *мм*. Ем-

[39] System puts data at arm's reach. Automatic setup stores information, delivers in seconds. «Iron Age», 1962, v. 189, № 2, стр. 72; L. H. Martin. The mechanized library. An experimental models. «Datamation», 1964, 10, № 9, стр. 32—36.

Рис. 86. Микрокопия документа и его поисковый образ на микрофильме ИПС «Rapid Selector»

и его поисковый образ (в кодированном виде) на микрофильме ИПС «Rapid Selector».

Плотность заполнения микрофильма зависит от применяемой кратности уменьшения. При кратности уменьшения 20 : 1 она составляет почти 140 *кадров/м*, а при кратности уменьшения 8 : 1 — примерно 20 *кадров/м*.

ИПС «Rapid Selector» состоит из трех следующих устройств: 1) установки для микрофильмирования; 2) модифицированного алфавитного перфоратора IBM-026 и 3) микрофильмового селектора. Ввод поискового образа документа в кодовое поле микрофильма осуществляется при помощи 80-колонной перфокарты, пробиваемой на модифицированном клавишном перфораторе IBM-026. Такая перфокарта подается в микрофильмирующую установку, в которой пробивки на перфокарте автоматически преобразуются в соответствующий двоичный фотокод, проецируемый на микрофильм.

В селекторах типа «Rapid Selector» используются бобины емкостью от 305 до 1800 *м*. На рис. 87 показана схема лентопротяжного механизма такого селектора. Из нее становятся ясными основные принципы работы этого селектора. Ввод поискового предписания в селектор производится при помощи 80-колонной перфокарты. Логические связи между дескрипторами поискового образа устанавливаются на коммутационной панели. При протяжке микрофильма специальный фотоэлектронный блок сравнивает с поисковым предписанием поисковые образы документов, микрокопии которых занесены на просматриваемый микрофильм. При совпадении срабатывает быстродействующая электромагнитная муфта, которая прижимает репродукционный микрофильм к ведущему барабану. Благодаря этому репродукционный микрофильм уже примерно через 3 *м/сек* протягивается с такой же скоростью, что и основной микрофильм. Когда копируемый кадр достигает положения экспозиции, его изображение через щелевую оптическую систему проецируется (в масштабе 1 : 1) на репродукционный микрофильм. Если поисковый образ следующего кадра не совпадает с введенным в селектор поисковым

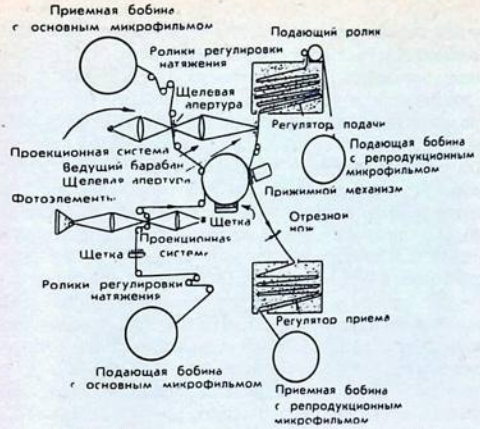



Рис. 87. Схема лентопротяжного механизма ИПС «Rapid Selector»

предписанием, то прижимная электромагнитная муфта выключается, репродукционный микрофильм отходит от ведущего барабана и останавливается. При следующем совпадении поискового образа документа с поисковым предписанием все перечисленные операции повторяются. Таким образом, в ИПС «Rapid Selector» рекопирование выдаваемых документов производится динамически, т. е. без остановки основного микрофильма. Это позволило применить в описываемой ИПС сравнительно высокую скорость протяжки основного микрофильма (1,8 *м/сек* и выше) и таким путем сократить среднее время информационного поиска.

В последней модели ИПС «Rapid Selector», которая ныне используется в Управлении кораблестроения и ремонта ВМС США, несколько модифицирован лентопротяжный механизм селектора: в нем для протяжки репродукционного микрофильма применен отдельный ведущий ролик и введены вакуумные магазины для компенсационных петель репродукционного микрофильма. Это позволило повысить скорость протяжки основного микрофильма до 2,7 *м/сек* и увеличить емкость основных бобин до 1800 *м*. При такой скорости протяжки микрофильма среднее время информационного поиска в одной бобине составляет около 12 мин. (включая проявление репродукционного микрофильма).

*FileSearch* [42]. ИПС «FileSearch» разработана фирмой FMA, Inc. (США) в 1959—1961 гг. и в настоящее время используется

[42] Information retrieval system cuts storage space, speeds access. «Electronics», 1961, v. 34, № 13, стр. 20—21; FileSearch retrieval unit. «National Micro-News», 1961, № 51, стр. 239; Automated system enhances BuShips

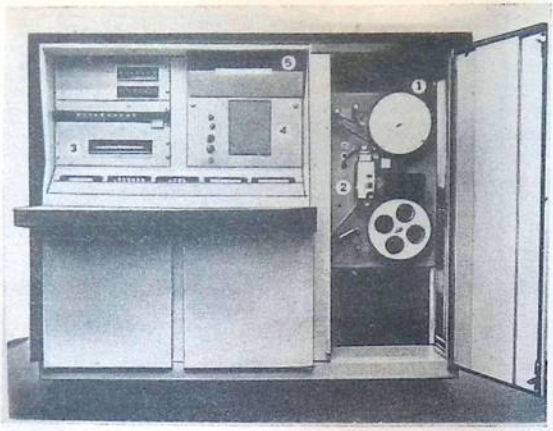



Рис. 89. Общий вид селектора ИПС «FileSearch»

*1* — механизм протяжки микрофильма; *2* — блок автоматической считки кодов; *3* — блок считки перфокарты с поисковым предписанием; *4* — выводной экран; *5* — блок изготовления увеличенных копий на бумаге

В случае необходимости может быть получена увеличенная копия найденного документа (продолжительность изготовления такой копии — от 5 до 10 сек.) или его микрокопия на репродукционном микрофильме (в масштабе 1 : 1). Так, используя селектор, можно создавать сводные микрофильмы, содержащие копии всех документов, в поисковые образы которых входит тот или иной общий дескриптор.

Поиск документальной информации при помощи ИПС «FileSearch» производится следующим образом. Сначала оператор переводит информационный запрос на применяемый в ИПС информационно-поисковый язык. Правильно сформулированное поисковое предписание при помощи флексорайтера пробивается на перфокарте, которая затем вводится в микрофильмовый селектор. Затем оператор отыскивает в хранилище бобину, содержащую микрокопии требуемых документов, и вставляет ее в селектор. Далее поиск производится автоматически.

При информационном поиске считанный с микрофильма код каждого дескриптора сравнивается с хранящимися в специальных регистрах кодами поискового предписания, и результаты сравнения запоминаются до того момента, когда фотоэлектронный блок считает первый код поискового образа следующего документа. После этого производится проверка на наличие между

соответствующими дескрипторами считанного поискового образа отношений И, ИЛИ,НЕ, БОЛЬШЕ и МЕНЬШЕ (эти отношения указываются в поисковом предписании). Если выполнены все условия смыслового соответствия, лентопротяжный механизм селектора останавливается и микрокопия первой страницы найденного документа автоматически устанавливается в положении визуального просмотра или копирования. После просмотра или копирования найденного документа все электронные схемы селектора переключаются в начальное состояние и поиск продолжается. Копирование найденных документов может производиться в увеличенном виде на специальную фотобумагу или автоматически на репродукционный микрофильм (в случае большого объема запрошенной документальной информации). Селектор может одновременно производить информационный поиск по шести поисковым предписаниям, если они имеют сравнительно несложную логическую структуру. Средняя продолжительность информационного поиска — от 2 до 3 мин. Функциональная блок-схема ИПС FileSearch показана на рис. 90.

*MIRACODE*[44]. ИПС MIRACODE[45] разработана фирмой Recordak Corp. (США) в 1962 г. В качестве носителя информации в этой ИПС используется 16-миллиметровый неперфорированный микрофильм (позитивный или негативный). Кодовые поля, в которых записываются поисковые образы документов, расположены перед кадрами, содержащими микрокопии страниц этих документов (рис. 91). Кодовое поле состоит из строк (рядов), каждая из которых содержит по 14 ячеек. В первых 12 ячейках строки записывается (в двоично-десятичном коде) трехразрядное десятичное число; 13-я ячейка (U) используется для контроля по четности. Кодовое поле может иметь различное число строк — до 15. Таким образом, максимальная полезная емкость кодового поля — 15 трехразрядных чисел[46]. Из сказанного следует также, что словарь информационно-поискового языка, используемого в ИПС MIRACODE, не может содержать более 999 дескрипторов.

Каждый кадр с микроизображением документа имеет свою метку, предназначенную для центровки кадра при его проеци-

---

[44] Recordak MIRACODE System. New York, Recordak Corp., 1963; M. M. Ruterbusch. Miracode — a new break-through in automated information retrieval. «Proceedings of the 12th Annual Meeting and Convention of the National Microfilm Association», Annapolis. National Microfilm Association, 1963, стр. 321—329; W. E. Perry, Jr. Microfilm systems. After the camera. «Administrative Management», 1963, v. 24, № 6, стр. 55; «The Recordak MIRACODE System. «Aslib Proceedings», 1964, v. 16, № 8, стр. 233; D. Jenkins. Automated document storage and retrieval. «Systems», 1964, v. 5, № 5, стр. 42, 44.

[45] MIRACODE=Microfilm Information Retrieval Access CODE — «Код доступа для поиска микрофильмированной информации».

[46] В первых вариантах ИПС MIRACODE максимальная емкость кодового поля составляла 24 двоичных единицы (шесть десятичных чисел).

Читально-копировальный аппарат «Lodestar Model PEK» предназначен для автоматического поиска кадров микрофильма по их шифрам (или поисковым образам), воспроизведения найденных кадров на просветном экране и изготовления на бумаге увеличенных копий с этих кадров. Скорость протяжки микрофильма при поиске — до 3 *м/сек*.

Особенностью читально-копировального аппарата типа «Lodestar» является то, что он снабжен системой автоматической заправки и перемотки микрофильма, рассчитанной на применение специальных закрытых кассет. Для поиска микрокопий нужных документов достаточно лишь вставить в аппарат типа «Lodestar» соответствующую кассету. Размер такой кассеты — 102×102××25 *мм*, ее емкость — 30,5 *м* микрофильма. Среднее время поиска в такой кассете — не более 8 сек., время изготовления копии на бумаге — 30 сек. Размер копии — 22×29 *см*. В комплект ИПС MIRACODE входит хранилище для микрофильмов емкостью 490 кассет (около 980 тыс. кадров). Аппарат «Lodestar Model PEK» снабжен консольной приставкой с клавишными пультами управления и ввода поисковых предписаний; в этой приставке смонтирован также быстродействующий логический блок.

Информационный поиск производится по одному трехразрядному десятичному числу (номеру или дескриптору). Возможны два варианта условий совпадения:

а) совпадающий код должен быть записан в строго определенной строке кодового поля;

б) совпадающий код может быть записан в любой из 15 возможных строк кодового поля.

Предусмотрена возможность подключения к аппарату «Lodestar Model PEK» до 14 клавиатур ввода. Это позволяет производить информационный поиск одновременно по 14 дескрипторам, коды которых в этом случае как бы образуют 42-разрядное десятичное число.

В настоящее время ИПС MIRACODE используется в Окриджской национальной лаборатории Комиссии по атомной энергии США и в Центре служб тыла Министерства обороны США (г. Бэттл Крик, штат Мичиган).

Необходимо отметить, что на базе читально-копировального аппарата типа «Lodestar» создан целый ряд различных устройств для адресного поиска микрофильмированных документов. Одним из простейших устройств такого типа является селектор *STORE* [47], используемый в Управлении разведки ВМС США с января 1961 г. В нем применен 16-миллиметровый позитивный микрофильм, на который с уменьшением 28 : 1 сфотографированы документы. Адресным шифром микрокопии первой страницы

[47] J. E. Willis. ONI's «STORE» provides rapid, reliable service. «Navy Management», 1963, v. 8, № 11, стр. 8—10; Electronisches Mikrofilm — Suchgerät Recordak ICK. «Reprographie», 1963, Bd. 3, № 7, стр. 165.

Рис. 92. Общий вид селектора «STORE» и схема расположения меток кадров на его микрофильме

каждого документа служит порядковый номер кадра, в котором находится эта микрокопия. Поэтому путем отсчета соответствующего числа кадров от начала микрофильма можно точно локализовать кадр с искомой микрокопией. На микрофильмах, используемых в селекторе STORE, каждый кадр снабжен прямоугольной меткой (рис. 92). Емкость бобины — 30,5 *м* (2500 кадров с микрокопиями). В качестве вводного устройства в селекторе STORE используется подключаемый к аппарату типа «Lodestar» блок управления с клавиатурой, который серийно выпускается фирмой Recordak Corp. На этой клавиатуре можно набирать четырехразрядные десятичные числа, являющиеся адресами документов на микрофильме. При включении лентопротяжного механизма фотоэлектронный блок аппарата типа «Lodestar» начинает считать метки кадров на микрофильме. Когда суммарное число таких меток сравнивается с набранным на клавиатуре, лентопротяжный механизм останавливается и найденный кадр

проецируется на экран аппарата. Как уже указывалось выше, аппарат типа «Lodestar» изготовляет увеличенные копии найденных микроизображений. Время поиска в бобине емкостью 30,5 *м* не более 10 сек.

*RADIR* [48]. ИПС RADIR создана фирмой Hallicrafters Co. (США) в 1963 г. и выполнена по двухконтурной схеме. В качестве носителя информации в обоих контурах этой ИПС используется 35-миллиметровый микрофильм. ИПС RADIR предназначена главным образом для хранения и поиска микрокопий технических чертежей и состоит из следующих устройств:

а) установки для покадрового микрофильмирования;

б) консольной установки для ввода на 35-миллиметровый микрофильм для поиска адресов документов по их поисковым образам;

в) микрофильмового селектора.

Установка для микрофильмирования снабжена кодирующим устройством, заносящим на микрофильм адресные шифры копируемых документов. Каждый такой шифр записывается в двоичном коде рядом с кадром, в котором содержится микрокопия соответствующего документа. С негативного микрофильма изготовляются позитивные копии (на пленке типа «кальфакс»), которые и используются в микрофильмовом селекторе ИПС. После контактного копирования негативный микрофильм поступает на архивное хранение. Соответствующим образом пронумерованные бобины с позитивными микрофильмами хранятся в стандартных фильмотеках. Общая емкость одной такой фильмотеки высотой 1,2 *м* — 300 тыс. кадров с микрокопиями технических чертежей формата E (432 × 533 *мм*).

Ввод информации в консольную установку ИПС RADIR, выполняющую функции $ЗУ_{вкт.}$, производится при помощи клавиатуры электрической пишущей машины. Каждая бобина микрофильма, используемого в этой установке, имеет емкость примерно 14 тыс. документов. Скорость просмотра поисковых образов — 700 *документов/сек*. Установка выдает в печатном виде поисковые образы документов и их адресные шифры. Бобины с микрофильмами для $ЗУ_{вкт.}$ ИПС создаются по предметному признаку.

Микрофильмовый селектор ($ЗУ_{нас.}$) ИПС RADIR предназначен для поиска микрокопий документов по их адресным шифрам. Ввод адресных шифров в селектор производится при помощи клавиатуры. Скорость поиска — 55 *кадров/сек*. Найденная микрокопия документа в увеличенном виде проецируется на экран

[48] Microfilm system is both compact and quick. «Factory», 1963, v. 121, № 8, стр. 141; Fast microfilm file system developed by Hallicrafters. «Electronic News», 1963, v. 8, N 384, стр. 29.

RADIR—Random Access Document Indexing and Retrieval — «Система индексирования и поиска документов с произвольным доступом».

Рис. 93. Общий вид селектора «CRIS»

селектора для ее визуального просмотра. Размер экрана — 180 × 230 *мм*. Возможно изготовление копий найденных микроизображений. Копирование производится контактным методом на 35-миллиметровую рулонную пленку типа «кальфакс». Время изготовления и проявления одной негативной копии (кадра) — 5 сек.

*CRIS* [49]. Селектор CRIS [50] разработан фирмой Retrieval Corp.[51] (США) в 1962 г. и выпускается фирмой Litton Systems, Inc. Впервые идея использования широкой рулонной фотопленки для хранения и адресного поиска документальной информации была высказана французом Ж. Себием в 1928 г. Позже эта идея нашла дальнейшее развитие в устройстве AMFIS [52], запатентованном американцем Э. Авакяном в 1952 г. (патентная заявка была подана в 1950 г.).

В качестве носителя информации в селекторе CRIS (рис. 93) используется рулонная диазопленка или пленка типа «кальфакс». Ширина пленки — 432 *мм*, длина — 61 *м* или 122 *м*. Позитивные микрокопии документов заносятся на эту пленку при помощи

[49] P. W. Larsen. CRIS... the Command Retrieval Information System. «Proceedings of the 11th Annual Meeting and Convention of the National Microfilm Association». Annapolis, National Microfilm Association, 1962, стр. 41—50; Information stored and retrieved on scrolls. «Business Automation», 1962, v. 8, № 3, стр. 48; W. E. Bushop. Information storage/retrieval. «Electronics», 1962, v. 35, № 26, стр. 55—57; E. P. Taubes. Microfilm as a systems tool. «National Micro-News», 1964, № 68, стр. 167.

[50] CRIS—Command Retrieval Information System — «Командная информационно-поисковая система».

[51] Фирма Information Retrieval Corp. является дочерним предприятием фирмы Information for Industry, Inc.

[52] AMFIS—Automatic Microfilm Information System — «Автоматическая система поиска микрофильмированной информации».

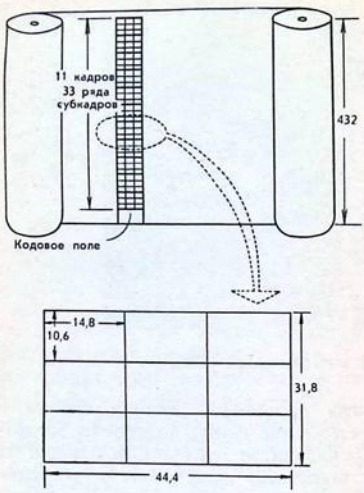



Рис. 94. Схема расположения микрокопий на рулонной пленке селектора «CRIS»

специальной установки для заполнения рулонов (Scroll Preparation Unit), входящей в комплект селектора CRIS. Эта установка снабжена шаговой кареткой и предназначена для рекопирования микроизображения документов с 16- и 35-миллиметровых галоидносеребряных негативных микрофильмов на рулонную пленку. По отношению к оригиналу кратность уменьшения микрокопий на рулонной пленке составляет 30 : 1. Установку для заполнения рулонов можно использовать для внесения дополнений в рулонную пленку типа «кальфакс». Для этой же цели предназначено устройство ввода дополнительных кадров (Add-a-Frame unit). В комплект селектора CRIS входит еще установка для контратипирования рулонной пленки.

Схема расположения позитивных микрокопий на рулонной пленке селектора CRIS показана на рис. 94. Микрокопии заносятся в кадры размером 31,8 × 44,4 *мм*. На рулонной пленке такие кадры расположены колонками по 12 кадров в колонке (в рулоне длиной 122 *м* содержится 2590 таких колонок). В колонке 11 кадров используется для микрокопий документов, а 12-й — для записи адресного шифра микрокопий, содержащихся в кадрах данной колонки. Кадры размером 31,8 × 44,4 *мм* используются для микрокопий карт и крупноформатных чертежей. Каждый кадр может быть разделен на 9 субкадров размером 10,6 × 14,8 *мм*. В одном субкадре размещаются микрокопии двух страниц формата 22 × 28 *см*. (Общая емкость рулона длиной 122 *м* — 511 тыс. таких страниц.) Пленка хранится в сменных кассетах размером 432 × 406 × 152 *мм*. Вес кассеты — около 11 *кг*.

Адресный шифр каждой микрокопии выражается семиразрядным десятичным числом. Этот шифр в двоичном коде записывается в 12-м кадре каждой колонки кадров. Ввод адресного шифра в селектор CRIS производится при помощи клавиатуры (возможен

также ввод на перфоленте). Для поиска нужного кадра вводится его семиразрядный адресный шифр, а также команда на проецирование всего кадра или первого субкадра этого кадра. Поиск колонки, в которой содержится кадр с искомой микрокопией, производится путем протяжки рулонной пленки из одной части кассеты в другую (в горизонтальной плоскости), а выбор в колонке нужного кадра — путем соответствующего сдвига всей кассеты относительно положения считывания (проецирования). При протяжке рулонной пленки фотоэлектронный блок сравнивает адресный шифр каждой колонки кадров с адресным шифром предшествующей колонки и таким образом контролирует правильность направления протяжки. Это избавляет от необходимости производить обратную перемотку пленки в подающую бобину кассеты перед началом поиска по каждому новому запросу. В селекторе CRIS поиск любого кадра начинается с той колонки, в которой находился кадр с микрокопией, запрошенной непосредственно перед искомым кадром. При совпадении адресного шифра, набранного на клавиатуре, с шифром, записанным в кодовом поле колонки, лентопротяжный механизм останавливается и найденный кадр (или его первый субкадр) проецируется на экран размером 457×610 *мм*. Скорость протяжки рулонной пленки при поиске до 4,6 *м/сек*. Среднее время поиска в рулоне длиной 122 *м* — 17 сек., а в рулоне длиной 61 *м* — до 10 сек.

В селекторе CRIS применены две оптические системы проецирования, что позволяет выводить на экран весь кадр или любой из субкадров этого кадра. Путем подачи соответствующих команд управления (при помощи клавиатуры) можно проецировать на экран поочередно все субкадры одного кадра, а также быстро переходить к следующему кадру. В случае необходимости найденная микрокопия может быть рекопирована контактным методом на отрезок пленки типа «кальфакс», вмонтированный в стандартную апертурную перфокарту. На одну апертурную перфокарту рекопируется один кадр или один субкадр с микрокопиями. Время рекопирования — не более 10 сек. Для чтения апертурных перфокарт и изготовления увеличенных отпечатков с вклеенных в них микрокопий можно использовать стандартную аппаратуру.

В настоящее время фирма Litton Systems, Inc. разрабатывает модифицированный вариант селектора CRIS, в который будет встроена выпускаемая этой фирмой фототелеграфная система типа Micropix-Printipix. Такая фототелеграфная система будет автоматически считывать найденные микроизображения, преобразовывать их в видеосигналы и передавать эти видеосигналы в установку Printipix для быстрой печати увеличенных изображений на бумагу. Возможна дистанционная передача таких видеосигналов. Производительность установки Printipix — 2,0—2,5 копии формата 22×28 *см/сек*. Для селектора CRIS разрабатывается также блок электрографического копирования. Очевидно, что

мый в блок сравнения. После этого включается лентопротяжный механизм, и фотоэлектронный блок начинает считывать с микрофильма адресные коды документов. При совпадении адресного кода с кодом запроса схемы управления останавливают лентопротяжный механизм и специальное устройство подрегулирования автоматически устанавливает кадр с микрокопией документа

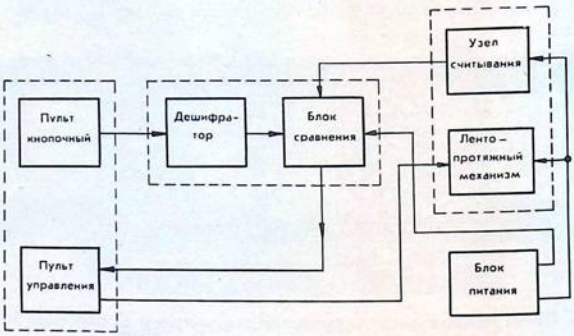



Рис. 97. Блок-схема селектора «Поиск-ОК»

точно в положение проецирования на просмотровый экран. Изображение документа на экране имеет размеры оригинала. Средняя продолжительность поиска микрокопий документов в бобине, содержащей 5 тыс. кадров (10 тыс. страниц), не превышает 4 мин.

В селекторе «Поиск-ОК» предусмотрена возможность получения электрографических копий найденных документов (в натуральную величину). Для этого в селектор вставляется кассета с предварительно заряженной пластиной из стандартного комплекта электрографического аппарата «ЭРА». На эту пластину проецируется увеличенное изображение микрокопии документа. Экспозицию при съемке контролирует реле времени, управляемое с пульта машины. Процессор для зарядки и проявления электрографической пластины, а также устройство для закрепления порошкового изображения на бумаге устанавливаются отдельно. На рис. 98 показана схема лентопротяжного механизма селектора «Поиск-ОК».

В настоящее время в ВИНИТИ изготовлен действующий макет микрофильмового селектора «Поиск-ДВ» (с динамическим выкопированием). Сохраняя в основном данные технической характеристики прежней модели, новый селектор (рис. 99) позволяет копировать необходимые кадры на 35-миллиметровый рольный микрофильм без остановки лентопротяжного механизма. Предусмотрено автоматическое управление селектором.

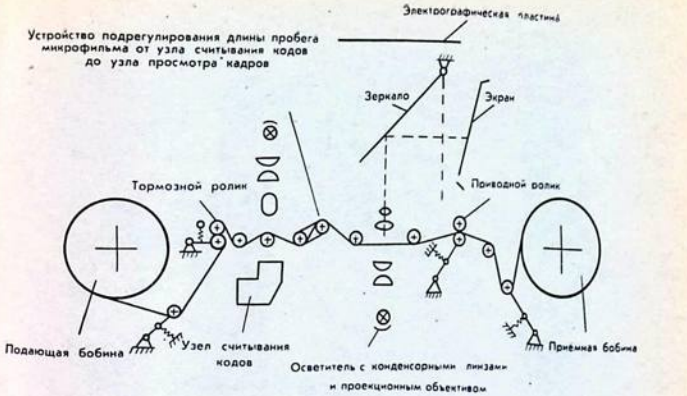



Рис. 98. Схема лентопротяжного механизма селектора «Поиск-ОК»

*FLIP* [55]. Микрофильмовый селектор FLIP [56] создан фирмой Benson-Lehner Corp. (США) в 1958 г. и предназначен для поиска микрокопий документов по их адресным шифрам. Носителям информации в селекторе FLIP является 16-миллиметровый перфорированный позитивный микрофильм. Кодовое поле, в котором записывается адресный шифр (номер) документа, расположено рядом с кадром, содержащим микрокопию этого документа; этот шифр записывается в виде комбинации зачерненных штрихов разной длины на прозрачном поле (рис. 100). Емкость кодового поля — 32 двоичных единицы. В селекторе применяются бобины емкостью 366 *м*. Плотность заполнения микрофильма — 197 *кадров/м*. Ввод адресного шифра запрашиваемого документа производится при помощи цифровой клавиатуры, расположенной на пульте управления селектора. Минимальная скорость протяжки микрофильма в селекторе — 1,5 *м/сек* (300 *кадров/сек*), максимальная — 3,0 *м/сек* (600 *кадров/сек*). Поскольку адресные номера на микрофильме расположены в порядке их возрастания, селектор сам определяет, в каком направлении необходимо перематывать микрофильм, чтобы быстрее вывести требуемый кадр в положение проецирования. При совпадении шифра, записан-

[55] G. Williams. FLIP: Film Library Instantaneous Presentation. «Library Resources and Technical Services», 1958, v. 2, № 4, стр. 278—281; High-speed FLIP. Automatic equipment for the searching of microfilm records. «Aircraft Production», 1960, стр. 216—217.

[56] FLIP—Film Library Instantaneous Presentation, что на русский язык можно перевести примерно так: «Мгновенная выдача материалов из микрофильмированной библиотеки».

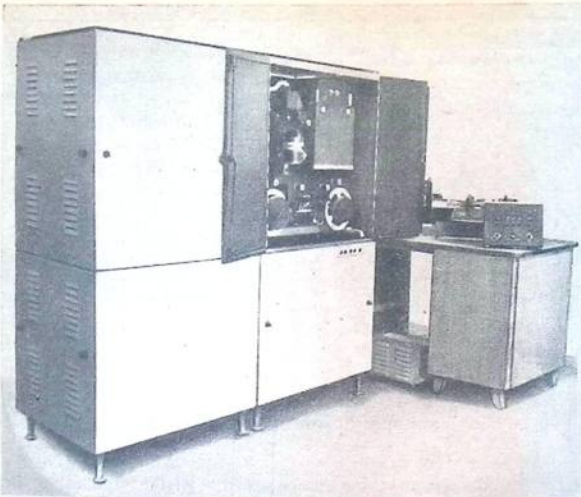


Рис. 99. Общий вид селектора «Поиск-ДВ»

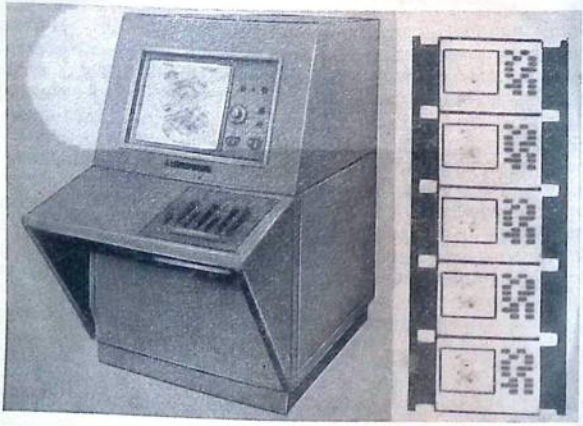


Рис. 100. Общий вид селектора ИПС «FLIP» и схема расположения микрокопий документов и их адресных шифров на микрофильме этого селектора

ного в кодовом поле микрофильма, с набранным на клавиатуре шифром лентопротяжный механизм селектора останавливается и кадр с микроизображением документа путем автоматической подрегулировки устанавливается в положение проецирования. Изображение документа проецируется на просветный экран с увеличением 1 : 50. Одновременно на специальном индикаторе в цифровой форме воспроизводится код этого кадра. Путем сравнения данного кода с набранным на клавиатуре шифром оператор может контролировать правильность работы селектора. Общее время пуска (разгона) и остановки лентопротяжного механизма селектора — 1,5 сек.

*Rapid Access Look-up System* [57]. Микрофильмовый селектор «Rapid Access Look-up System» [«Каталожная система с быстрым доступом»] создан фирмой Ferranti-Packard Electric Ltd. (Канада) в 1960 г. Он предназначен для поиска микрокопий документов по их адресным шифрам. Носителем информации в этом селекторе является 16-миллиметровый позитивный микрофильм. В кодовом поле каждого кадра этого микрофильма записывается шифр, идентифицирующий страницу каталога, микрокопия которой содержится в этом кадре. На каждой странице каталога содержатся сведения о пяти изделиях. Ввод адресных шифров в селектор производится при помощи клавиатуры, которая может быть использована также для пробивки перфокарт. Скорость протяжки микрофильма в селекторе — свыше 1 *м/сек*. При нахождении нужного кадра лентопротяжный механизм останавливается и на экран селектора проецируется изображение страницы каталога, на которой содержатся сведения об интересующем изделии. Благодаря применению такого селектора среднее время поиска в каталоге объемом 500 страниц (кадров) составляет всего около 3,5 сек. В настоящее время для этого селектора разрабатывается блок увеличения и печати на бумагу найденных микрокопий. Семь селекторов такого типа используется фирмой Ellicott Drug Co. (г. Буффало), один — Международной ассоциацией машинистов (г. Вашингтон) и еще один — страховой компанией Home Insurance Co. (Нью-Йорк).

Существуют или создаются и другие типы и варианты микрофильмовых селекторов (например, FOSDIC-2, Téleccetron и др.), но они по конструкции мало чем отличаются от уже описанных устройств.

В 1959 г. фирма International Business Machines Corp. (США) создала *экспериментальный микрофильмовый селектор*, в котором поисковые образы документов записываются на магнитной

[57] Microfilm search library. «Automatic Data Processing», 1961, v. 3, № 12, стр. 37—38; Cataloguing on film. «Machinery Lloyd (Overseas Edition)», 1962, № 8, стр. 52; New electronic systems aid office automation. Rapid Access Look-up System. «Process Control and Automation», 1962, v. 9, № 4, стр. 172—174.

| Название | Тип поиска | Страна | Год создания | Ширина носителя, мм | Кратность уменьшения | Средняя плотность заполнения микрофильма, кадры/м | Емкость кодового поля, двоичные единицы |
|---|---|---|---|---|---|---|---|
| Rapid Selector-1 | С | США | 1949 | п35П | | 118 | 228 |
| Rapid Selector-2 | С | США | 1959 | п35П | 8*/20* | 10/138 | 276Х |
| Rapid Selector-2a | С | США | 1962 | п35П | 2,5 ÷ 20* | | 315Х |
| FOSDIC-2 | С | США | 1957 | п16 | $\frac{24*}{43,5*}$ | 433 | 960 |
| FLIP | А | США | 1958 | п16П | 50* | 197 | 32 |
| Photoelectric Information Selector | С | Австралия | 1959 | п16Н | | 130 | 150—175 |
| Rapic Access Look-up System | А | Канада | 1960 | 16П | | | |
| FileSearch | С | США | 1961 | п35 | 25* | 105 | 392 |
| Teleccetron | С+А | Франция | 1961 | 16Н | 20*/40* | | 9 к. с. |
| Lodestar Model PEK (с блоком ввода адресов) | А | США | 1961 | н16П | 24* | 115 | 1 |
| MIRACODE | С | США | 1962 | н16Н/П | 12 ÷ 28* | 66 | 210 |
| Поиск-ОК | А | СССР | 1962 | н35П | 15* | 19 | 30 |
| CRIS | А | США | 1962 | н432П | 30* | 2094 | 28 |
| RADIR | С+А | США | 1963 | 35П | | | |
| 70 mm Selector | | США | | н70 | 4 ÷ 14* | 10 | 960 |

* Кратность уменьшения по ширине документа (перфокарты) 24 : 1, а по длине — 43,5 : 1.
П — позитивный; п — перфорированный; С — поиск по содержанию.

Таблица 30

**микрофильмовых селекторов**

| Скорость протяжки микрофильма | | Емкость бобины, м | Среднее время поиска | Способ вывода найденных документов | | | | Ориентировочная стоимость, тыс. долл. | Примечание |
|---|---|---|---|---|---|---|---|---|---|
| м/сек | кадры/сек | | | экран | репрогукционный микрофильм | апертурные перфокарты | увеличенные копии на бумаге | | |
| 1,5 | 180 | 610 | 7 мин | — | + | — | — | 75 | |
| 1,8 | 36 | 610 | 6 мин | — | + | — | — | 80 | |
| 2,7 | | 1800 | 12 мин | — | + | — | — | | |
| 0,15 | 67 | 30,5 | 3,4 мин | — | — | — | — | 112 | Вывод на перфокартах |
| 1,5/3,0 | 300/600 | 336 | 2 мин | + | — | — | — | 40—50 | |
| | 3 | 122 | 87 мин | — | — | — | — | | |
| 1 | | 1000 кадров | 7 сек | + | — | — | — | 17 | Вывод номеров документов в печатном виде |
| 1 | 106 | 305 | 2—3 мин | + | + | — | + | 143—157 | |
| 1 | 50 | 30,5 | 30 сек | + | — | — | — | 9 | |
| 3 | 345 | 30,5 | 5 сек | + | — | — | + | 4,25 | |
| 3 | 198 | 30,5 | 8 сек | + | — | — | + | 30 | |
| 0,67 | 13 | 250 ÷ 270 | 4 мин | + | — | — | + | | |
| 4,6 | 9632 | 61/122 | 10/17 сек | + | — | + | — | 33,5 | |
| | 700C / 55A | | | + | — | + | — | 35 | Двухконтурная ИПС |
| 0,5 | 5 | 3P5 | 10 мин | + | — | — | + | 80 | В стадии разработки |

А — поиск по адресу; к. с. — ключевые слова; Н — негативный; н — неперфорированный;

министерства ВВС США. В качестве носителя информации в ней используется магнитная лента шириной 25,4 *мм*. Емкость бобины с магнитной лентой — 731,5 *м*. На магнитной ленте записываются поисковые образы документов, а также их адресные шифры (с некоторой уточняющей информацией). Объем поискового образа — до 140 знаков. В одной бобине магнитной ленты может быть записана информация примерно 70 тыс. документов. Ввод в ИПС поисковых образов документов с их адресными шифрами, а также поисковых предписаний производится на перфоленте. Логические отношения между дескрипторами поискового предписания вводятся в систему с коммутационной панели. ИПС типа «CCC Index Searcher» может производить информационный поиск одновременно по нескольким поисковым предписаниям. Время поиска в бобине емкостью 731,5 *м* (70 тыс. документов) — около 4,5 мин. Результаты информационного поиска выдаются в печатном виде (в качестве устройства ввода-вывода в ИПС типа «CCC Index Searcher» используется флексорайтер). ИПС выдает адресные шифры искомых документов, а также номера информационных запросов, по которым производился информационный поиск. Однако она может выдавать вместе с адресными шифрами документов и их поисковые образы. Стоимость ИПС типа «CCC Index Searcher» — около 120 тыс. долл.

*GE-250 Information Searching Selector*. ИПС типа «GE-250 Information Searching Selector» была создана фирмой General Electric Co. (США) в 1959 г. для университета Western Reserve (г. Кливленд, штат Огайо). В ИПС применяются магнитные ленты шириной 12,7 *мм* и длиной 731,5 или 1097 *м*. На каждой такой ленте записываются поисковые образы и адресные шифры соответственно 13 тыс. или 19 тыс. документов. Объем поискового образа документа с его адресным шифром — до 300 буквенно-цифровых знаков. В качестве устройства ввода-вывода в ИПС типа «GE-250 Information Searching Selector» используется флексорайтер. Информационный поиск можно производить одновременно по нескольким предписаниям (до 10). Время поиска в бобине емкостью 731,5 *м* — не менее 16 мин., а в бобине емкостью 1097 *м* — 24 мин.

*Search Comparator*. ИПС типа «Search Comparator» создана фирмой General Electric Co. в 1961 г. и является уменьшенным вариантом ИПС типа «GE-250 Information Searching Selector». Система позволяет информационный поиск одновременно по нескольким запросам. Время поиска в одной бобине — менее 5 мин. Стоимость ИПС — около 50 тыс. долл.

*Heatwole H-44*. ИПС типа «Heatwole H-44» создана американской фирмой Heatwole Associates (Herner Inc.) в 1961 г. В качестве носителя информации в ней используется магнитная лента, на которой записываются поисковые образы документов вместе с их адресными шифрами. Поисковый образ документа может

состоять из любого числа дескрипторов, записываемых в виде трех-, четырех- и пятиразрядных кодов, а адресный шифр документа может иметь до девяти десятичных разрядов. Емкость одной бобины с магнитной лентой — около 37,5 тыс. документов (если средний объем поискового образа — четыре четырехразрядных кода, а средний адресный шифр — пятиразрядное десятичное число). Ввод поисковых предписаний в ИПС производится с панели управления при помощи системы переключателей. Поиск производится по четырем дескрипторам, которые могут быть связаны логическими отношениями И, ИЛИ и НЕ. Скорость сканирования магнитной ленты при поиске — около 1000 документов в мин. Стоимость ИПС типа «Heatwole H-44» — 9,5 тыс. долл.

*RCA Video File*. Экспериментальная ИПС типа «RCA Video File» разработана фирмой Radio Corporation of America в 1962 г. В качестве носителя информации в ней используется магнитная лента шириной 50,8 *мм*, на которой записываются изображения документов и их поисковые образы или адресные шифры. Емкость одной бобины магнитной ленты — 1829 *м*. На магнитной ленте такой длины можно записать до 36 тыс. страниц документов. Время поиска документов в одной бобине — около 5 мин. Скорость видеомагнитной записи — считывания изображений документов — примерно 100 *стр/мин*.

Изображения найденных документов проецируются на телевизионный экран, фотографируются или воспроизводятся в натуральном размере на полупроводниковой бумаге со скоростью 1—2 *стр/сек*. Стоимость ИПС типа «RCA Video File» — от 0,5 до 1,5 млн. долл.

*Videofile* [59]. ИПС «Videofile» создана фирмой Ampex Corp. (США) в 1958—1964 гг. и предназначена для Центра по космическим полетам им. Джона Маршалла (г. Хантсвилл, штат Алабама). В ней будут храниться (в видеомагнитной записи) научно-технические отчеты о результатах испытаний на надежность деталей и узлов ракет. Предполагается, что этой ИПС будут пользоваться все научно-исследовательские и проектно-конструкторские центры Национального управления по аэронавтике и исследованию космического пространства США, а также его главные подрядчики. Все они должны быть оснащены соответ-

---

[59] R. A. Minor. Television — tomorrow's answer. «Reproduction Engineer», 1963, v. 7, № 1, стр. 8—10, 20; R. A. Minor. Magnetic tape: a new microreproduction tool. «Proceedings of the 12th Annual Meeting and Convention of the National Microfilm Association», Annapolis, National Microfilm Association, 1963, стр. 107—116; Television filing system «The Office», 1964, v. 60, № 4, стр. 145; Micro-storage of documents replaces file folders. — «Journal of Machine Accounting», 1964, v. 15, № 12, стр. 22—23; Recording data on video tape. «Data Processing (Engl.)», 1965, v. 7, № 2, стр. 62—64; The Videofile system. «Datamation», 1965, № 1, стр. 57.

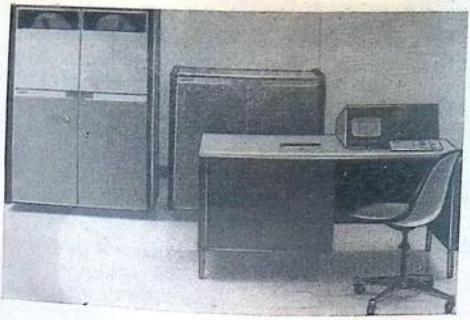


Рис. 104. Общий вид ИПС «Videofile»

документов. В конце дня полученный список используется для перезаписи изображений документов с временной видеомагнитной ленты на постоянную.

Поиск требуемых документов может производиться несколькими способами. В простейшем случае для этого достаточно набрать на клавиатуре блока индексирования адресный шифр искомого документа, и ИПС «Videofile» воспроизведет изображение этого документа на экране монитора или выдаст его электрографическую копию. Скорость адресного поиска в ИПС «Videofile» не менее 500 *кадров/сек*. Общий вид ИПС «Videofile» показан на рис. 104.

Основное назначение ИПС «Videofile» — поиск документов по их адресным шифрам и выдача копий этих документов (функция $ЗУ_{нас.}$). Но в Центре по космическим полетам им. Джона Маршалла ее планируется использовать в соединении с соответствующей ЭЦВМ, которая будет выполнять функции $ЗУ_{акт.}$ в двухконтурной документальной ИПС. Предполагаемая стоимость ИПС «Videofile» — от 0,2 до 1,0 млн. долл.

Глава

8

# КОПИРОВАНИЕ И РАЗМНОЖЕНИЕ ДОКУМЕНТОВ (репрография)

В течение сотен лет со времени изобретения И. Гутенберга типографский процесс был единственным способом многотиражного размножения текстовых документов. Лишь в последнее столетие в качестве средства копирования документов стала применяться изобретенная в 30-х годах прошлого века фотография. Потребность в быстрых и дешевых способах и средствах копирования, возникшая несколько десятилетий назад в сфере делопроизводства и организации труда, привела к разработке многочисленных процессов копирования и размножения: рефлексных, диазографических, термографических, электрографических и др.

Быстрое распространение и широкое применение новых методов и средств копирования и размножения вызвали к жизни обширную литературу: описания процессов копирования, принципов действия и устройства новых аппаратов, соображения о целесообразности и экономичности их применения для разных целей и т. п. Постепенно стала складываться прикладная дисциплина о различных средствах копирования и размножения документов. В начале 50-х годов для обозначения этой новой области был предложен термин «репрография». В настоящее время этот термин можно считать общепризнанным. С 1961 г. в ФРГ издается на немецком языке журнал «Reprographie», а в 1963 и 1967 гг. состоялись международные конгрессы по репрографии.

Репрография объединяет представителей самых различных специальностей: инженеров в области техники копирования и размножения документов, организаторов народного хозяйства, науки и культуры, экономистов, специалистов в области научной информации, библиотекарей, библиографов и др.

Что касается информатики, то она имеет особые, специфические интересы в области репрографии. Создание современных информационно-поисковых устройств, описанных в предыдущей

ский процесс, студент медицинского факультета Льежского университета Альбрехт Брейер начал производить эксперименты по использованию галоидного серебра для покрытия бумаги с целью получений копий со страниц печатной книги.

В настоящее время при помощи обычных фотографических процессов изготавливаются копии высокого качества, выдерживающие длительное архивное хранение. Однако эти процессы слишком дороги, требуют больших затрат времени и специально оборудованного затемненного помещения. В научно-информационной практике обычные фотографические процессы применяются главным образом для изготовления копий полутоновых иллюстраций с целью их последующего тиражного размножения (фотостатное копирование), а также для изготовления копий с уменьшением (микрокопирование) или с увеличением (репродукционное копирование).

Изготовление сильно уменьшенных копий документов называется *микрофотографированием*, или *микрофильмированием*. По сравнению с оригиналами микрокопии имеют следующие преимущества: уменьшение объема оригинальных документов при хранении не менее чем на 95%, стандартизация размеров низкая стоимость.

При помощи микрофотографирования можно оперативно выполнять копирование документов, их малотиражное размножение, также выпуск периодических и непериодических микроизданий. Микрофотографирование включает два процесса: изготовление ативов микрокопий и печатание с них позитивов. Негативы рокопий служат натуральную величину. Позитивы микрокопий ис-ользуются непосредственно для чтения на читальных аппаратах. ликрофильмирование документов производится на аппаратах ической (покадровой) и динамической (щелевой) съемки. При ровой съемке изображение оригинала проецируется на непо-ную фотопленку, происходит одновременное экспонирование кадра. Принцип работы аппаратов щелевого типа заключается ографировании движущегося оригинала на синхронно дви-ося пленку в момент прохождения оригиналом съемочной ы. В сравнении с покадровыми, щелевые микрофильмирующие ие аппараты более высокую скорость съемки и значительно имеют габариты. На основе принципа щелевой съемки разра-ы для одновременного двухстороннего микро-листовых документов при скорости их переме-*м/мин*. В аппаратах щелевой съемки может ис-самонаклад, обеспечивающий высокую скорость ирования однотипных оригиналов.

экспонированных микрофильмов производится на ройствах для проявления фиксирования и сушки на автоматических проявочных машинах. В про-

явочной машине рулонная пленка непрерывно проходит через все стадии химической обработки: проявление, фиксирование, промывку и сушку.

Копирование с микрофильмов производится контактным и проекционным способами. Проекционный способ применяется при изготовлении копий измененного масштаба. При копировании контактным способом негативный микрофильм накладывается эмульсионным слоем на такой же слой неэкспонированного микрофильма и освещается со стороны негатива. В таком виде обе пленки протягиваются мимо печатающего (освещаемого) окна специальных копировальных аппаратов, работающих со скоростью более 100 *м/час*.

В последнее время все более широкое распространение получают микрофотокопии, которые выполняются в виде *микрокарт*, имеющих прозрачную или непрозрачную подложку.

Чтение и просмотр микрофотокопий осуществляется при помощи специальных читальных аппаратов, которые могут иметь отражающий или просветный экран. Выпускается большое количество стационарных, настольных и карманных читальных аппаратов, которые обеспечивают увеличение от 5 до 40 раз. В большинстве из них микрофотокопии читаются в проходящем свете.

Для получения увеличенных копий с микрофильмов могут быть использованы как обычные фотоувеличители, так и специальные электрографические аппараты. Кроме того, все большее распространение получают читально-копировальные аппараты, на которых после чтения или просмотра микрокадра можно с рулонной пленки или с микрокарты сразу же получить на специальной бумаге увеличенную копию этого микрокадра.

### *Рефлексные процессы*

Для того чтобы устранить трудности, связанные с обычными фотографическими процессами, сделать копирование с использованием галоидного серебра более быстрым и не требующим затемненного помещения, были разработаны рефлексные процессы, в основу которых был положен метод контактного копирования, когда экспонирование фотобумаги производится светом, отражающимся от поверхности копируемого оригинала. Из них наибольшее распространение получили стабилизационный, диффузионный и матричный процессы.

*Стабилизационный* процесс разработан для того, чтобы исключить необходимость удаления непрореагировавшего галоидного серебра с обычной фотографической бумаги, после того как эта бумага была экспонирована и проявлена. При обычном фотографическом процессе непрореагировавшее галоидное серебро удаляют при помощи растворителя (процесс фиксирования). В качестве такого растворителя обычно используется тиосульфат

необходимо применение непрямого рефлексного метода. Этот метод заключается в том, что прямопозитивная фотобумага накладывается на копируемый оригинал не эмульсионным слоем, а подложкой. Освещение производится со стороны эмульсионного слоя. При таком методе экспонирования свет, отражающийся от белых участков оригинала, попадает на эмульсионный слой фотобумаги лишь после прохождения сквозь ее подложку. Для обеспечения четкости изображения между источником света и эмульсионным слоем обычно устанавливают полупрозрачный экран из пластмассы.

*Диффузионный процесс*, или процесс переноса проявленного изображения методом диффузии, требует применения двух типов специальной бумаги — негативной и позитивной. Негативная бумага покрыта желатиновой эмульсией, содержащей галоидное серебро и другие вещества. Эта эмульсия чувствительна к свету. Позитивная бумага покрыта слоем коллоидального серебра и нечувствительна к свету.

Для экспонирования негативная бумага накладывается эмульсионным слоем на лицевую сторону копируемого оригинала. Освещение производится с обратной стороны негативной бумаги. При этом выбираются такая чувствительность эмульсионного слоя и такое время экспозиции, чтобы свет, однократно проходящий через этот слой, не оказывал на него никакого действия. Скрытое изображение образуется лишь на тех участках галоидно-серебряного слоя, которые подверглись воздействию как прямого, так и отраженного света. Эти участки соответствуют таким местам копируемого оригинала, в которых нет изображения. После экспонирования негативную бумагу отделяют от оригинала, накладывают лицевой стороной на лицевую сторону позитивной бумаги в таком виде погружают их в раствор, содержащий проявитель (например, метол и гидрохинон) и растворитель галоидного серебра (например, тиосульфат натрия).

Проявитель восстанавливает галоидное серебро экспонированных эмульсионных микрокристаллов в металлическое серебро. Растворитель происходит в местах скрытого изображения негатива. Растворитель галоидного серебра взаимодействует с неэкспонированным галоидным серебром на участках, соответствующих изображению на оригинале, и образует растворимые комплексные соединения серебра. Эти комплексные соединения благодаря диффузии переносятся с негативной бумаги на позитивную, которая прижата к эмульсионной поверхности негатива.

Эмульсионный слой позитивной бумаги содержит реагент, который в присутствии проявителя выделяет из комплексного соединения металлическое серебро. В качестве такого реагента применяется коллоидальное серебро. Для этой цели можно использовать сернистое серебро, коллоидальную ртуть и также ее органические соединения.

Для обеспечения необходимого контакта между эмульсионными слоями негативной и позитивной фотобумаги, а также для отжима раствора сложенные листы негативной и позитивной фотобумаги пропускаются через специальные валики и после этого отделяются один от другого. Эта операция обычно выполняется вручную. Позитивный лист, на котором имеется черно-белое изображение оригинала, затем в течение нескольких минут сушится на воздухе. Обычно с одного негатива изготовляют только одну позитивную копию (рис. 106).

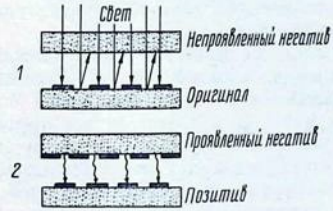



Рис. 106. Схема диффузионного процесса

*1* — от светлых участков оригинала свет отражается на светочувствительный слой, образуя негативное изображение; *2* — методом диффузии неэкспонированное галоидное серебро переносится на позитив и преобразуется в металлическое серебро

Так как позитивная копия получается во время проявления негатива, то этот способ является очень быстрым. Раньше продукты фиксирования, остававшиеся на бумаге после ее обработки, приводили к появлению на копиях коричневых пятен. Однако вследствие применения сильных щелочных растворов эта реакция замедляется, что позволяет хранить копии при нормальной влажности и температуре в течение нескольких лет.

Другой недостаток, который является следствием появления продуктов окисления проявителя, заключается в пожелтении копий. Для предотвращения этого в проявитель добавляются соответствующие отбеливающие вещества.

*Матричный процесс* (его также называют процессом переноса изображения при помощи желатина или методом «верифакс») основан на использовании матрицы — листа светочувствительной бумаги, которая покрыта слоем незадубленного желатина, содержащего галоидное серебро, дубящий проявитель и цветообразующее вещество. Экспонирование матрицы производится рефлексным способом.

Экспонированная матрица погружается в щелочной активатор, содержащий углекислый натрий и мягчитель желатина,

который в дальнейшем способствует осуществлению операции переноса изображения.

Под воздействием активатора желатиновый слой матрицы набухает, а дубящий проявитель, в качестве которого может использоваться гидрохинон или пирокатехин, химически восстанавливает галоидное серебро в металлическое на экспонированных участках матрицы. При этом дубящий проявитель окисляется кислородом воздуха, растворенным в активаторе. Продукт окисления проявителя вызывает затвердение желатина на тех участках, которые были подвергнуты воздействию света. Остальные участки желатинового слоя, соответствующие элементам копируемого изображения, не затвердевают.

Во время проявления некоторая часть цветообразующего вещества окисляется и, соединяясь с неокисленным веществом, образует в желатиновом слое нерастворимый синий краситель. Этот краситель может быть перенесен на другой лист бумаги только путем физического переноса желатина на тех участках слоя, где он не затвердел и стал нерастворимым.

Избыток активатора отжимается из набухшей матрицы путем прижатия ее к желатиновому слою листа бумаги, обработанной тиомочевиной. Копии можно получать и на обычной необработанной бумаге. Однако ее предварительная обработка тиомочевиной обеспечивает получение более темного и долговечного изображения.

Когда к лицевой стороне матрицы прижимается лист копировальной бумаги, то на поверхность этой бумаги переносится тонкий слой желатина, причем это происходит лишь на тех участках, где слой желатина набух и не затвердел. Перенесенный желатин содержит не только синий краситель, но и галоидное серебро, дубящий проявитель и небольшое количество металлического серебра.

Краситель непосредственно и образует копируемое изображение. Следует отметить, что этот краситель уже через несколько недель выцветает. Но если копировальная бумага была предварительно обработана тиомочевиной, то выцветание красителя в значительной степени компенсируется постепенным преобразованием перенесенного галоидного серебра в металлическое и в сульфид серебра.

После пропускания матрицы с наложенной на нее копировальной бумагой между сжимными валиками от матрицы вручную отделяется готовая слегка влажная позитивная копия (рис. 107). С одной матрицы можно получить до 10 копий. Для этого перед операцией переноса изображения матрицу нужно каждый раз погружать в раствор активатора.

Светочувствительность желатинового слоя, которым покрыта матрица, достаточно низка. Поэтому копирование можно проводить в помещении с флуоресцентным освещением (величина осве-

щенности — порядка 430 *лк*). Однако этот слой очень чувствителен к желтому цвету. Поэтому при использовании матричного процесса трудно копировать желтые, оранжевые и зеленые оригиналы. Этот недостаток присущ большинству процессов, основанных на рефлексном способе экспонирования. Кроме того, матричный процесс не обеспечивает — без растровой маски —

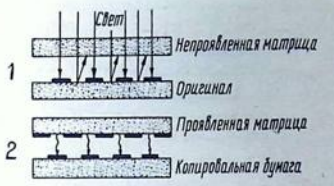



Рис. 107. Схема матричного процесса («верифакс»)

*1* — от светлых участков оригинала свет отражается на желатиновый слой матрицы; *2* — незадубленные участки красящего желатинового слоя переносятся на копировальную бумагу

хорошего качества копирования полутоновых изображений: светлые тона получаются слишком белыми, а серые — слишком темными. Так как температура активатора является критической характеристикой (его температура должна поддерживаться на уровне 27—28° С), то копировальный аппарат необходимо включать примерно за час до начала работы на нем. Еще один недостаток матричного процесса заключается в том, что он может хорошо выполняться лишь подготовленным оператором.

Выше уже говорилось, что бумага, на которой изготовляются копии, не имеет специального покрытия. Поэтому на ней можно делать любые пометки. Копии можно изготовлять с обеих сторон бумаги. При этом процессе бумага меньше скручивается, чем при любом другом процессе, основанном на рефлексном способе экспонирования.

### *Диазографические процессы*

Диазография, называемая также *светокопированием* и *диазотипией*, применяется главным образом для снятия копий с технических чертежей. Диазография является самым дешевым способом копирования: стоимость диазографической копии всегда в четыре-пять раз ниже стоимости копии, полученной любым другим способом. Однако серьезным недостатком диазографии является то, что копируемый оригинал должен быть выполнен на одной стороне прозрачного или полупрозрачного материала.

глазом как белые. Там же, куда не попали УФ-лучи, фотослой после проявления остался прозрачным. Сквозь такие прозрачные участки видна черная подложка, что воспринимается как элементы изображения.

Фототермография, или метод «кальвар», не требует мокрой обработки экспонированных фотослоев и темного помещения, а температура проявления не настолько высока, чтобы создать какие-либо серьезные проблемы. Стоимость бумаги и пленки типа «кальфакс» не превышает стоимости галоидносеребряных материалов, а долговечность изготовленных на такой бумаге или пленке копий исключительно высока.

### *Термографические процессы*

Из термографических процессов наиболее распространенным является процесс *термографии без переноса изображения* или «термофакс», разработанный фирмой «Minnesota Mining and Manufacturing Co.» (США) в 1950 г. Этот процесс основан на использовании термореактивной бумаги, которая содержит вещества, образующие при нагревании окрашенное соединение. Для такой реакции необходимо соединение трехвалентного железа (например, стеарат или миристат трехвалентного железа), а также соединение фенола (например, пирогаллол, галловая или дубильная кислота). Эти вещества содержатся в каком-либо связующем веществе (например, в поливинил бутирале). Преждевременная реакция между этими веществами предотвращается тем, что соединение трехвалентного железа применяется в воскообразном состоянии. Реакция с образованием окрашенного соединения начинается лишь в тех случаях, когда соединение трехвалентного железа переходит в расплавленное состояние.

При копировании лист термореактивной бумаги накладывается обратной стороной на поверхность копируемого оригинала и подвергается ИК-облучению (с лицевой стороны термореактивной бумаги). ИК-лучи проходят сквозь термореактивную бумагу и, поглощаясь темными участками копируемого изображения, нагревают контактирующие с ними участки термореактивной бумаги. Но это происходит лишь тогда, когда копируемое изображение выполнено красителем, содержащим газовую сажу или соединение какого-либо металла. В результате такого нагрева соединение трехвалентного железа приходит в расплавленное состояние и, вступая в реакцию с соединением фенола, образует на термореактивной бумаге позитивную копию оригинала (рис. 109).

Первоначально для процесса «термофакс» был разработан другой тип термореактивной бумаги. Эта бумага представляла

собой прозрачную для ИК-лучей темную подложку, покрытую кристаллическим карнаубским воском или стеаратами. Это покрытие непрозрачно для видимой части спектра и имеет белый цвет. На термореактивной бумаге такого типа копия изображения образуется лишь за счет изменения физического состояния покрытия: на участках бумаги, контактирующих с элементами копируемого изображения, восковое покрытие расплавляется и, перейдя в прозрачное состояние, делает видимой темную подложку бумаги. Однако такая термореактивная бумага выпускалась всего два года, так как она не обеспечивала достаточной четкости копий и была чувствительной к царапинам и т. п.

Рис. 109. Схема термографического процесса

Тепло, образующееся при поглощении инфракрасных лучей темными участками оригинала, воспринимается термореактивной бумагой, на которой формируется изображение

Выше уже отмечалось, что процесс «термофакс» не позволяет копировать изображения, выполненные красителями растительного происхождения. Для устранения этого недостатка была разработана специальная приставка к термокопировальному аппарату, при помощи которой сначала с цветных оригиналов (в том числе сброшюрованных) изготовляется промежуточная матрица. Далее эта матрица используется в качестве оригинала для изготовления копий на обычной термореактивной бумаге. Такая матрица представляет собой лист прозрачного материала, который покрыт веществом, чувствительным к двум разным участкам светового спектра (отсюда происходит и фирменное название этого процесса — дуаль-спектровый). Экспонирование матрицы производится рефлексным способом. Лучи видимого света, отражаясь от пробельных участков копируемого оригинала, делают нечувствительными к свету соответствующие участки матрицы. Неразложившееся светочувствительное вещество образует на матрице скрытое изображение оригинала. Затем на экспонированную матрицу накладывается лист копировальной бумаги, предварительно подвергнутой специальной химической обработке. При нагревании матрицы с контактирующей с нею копировальной бумагой до 100° С вещество, содержащееся в копировальной

бумаге, вступает в реакцию с неразложившимся светочувствительным веществом матрицы, образуя на копировальной бумаге позитивное изображение оригинала. Для такого нагрева экспонированная матрица, выполняющая функцию оригинала, вместе с наложенным на нее листом копировальной бумаги пропускается через термокопировальный аппарат, предназначенный для копирования методом «термофакс».

*Термографическое копирование с переносом изображения*. Из термографических процессов такого типа наиболее известны методы «мастерфакс», «эктафакс» и метод Эйхнера. При использовании каждого из этих методов копия может быть получена на любой бумаге. Для копирования листовой оригинал вместе с обычной бумагой и матрицей, покрытой теплочувствительным красителем, вкладывается в обложку из шелковой сетки и подвергается ИК-облучению. После экспонирования эти три листа пропускаются между пресс-валиками, в результате чего расплавленный теплом краситель переходит на бумагу. Такой процесс пригоден для копирования изображений, выполняемых красителем с металлической основой.

При термографическом копировании с переносом изображения изготовление одной копии расходуется больше материалов, трица при использовании процесса «термофакс». Однако каждая этому при изготовлении до шести копий с одной матрицы термо-фические процессы с переносом изображения могут обеспе-ать получение даже более дешевых копий, чем процесс мофакс».

Основным достоинством термографических процессов с пере-ным изображения является то, что они позволяют получать ные матрицы для спиртовой печати. Так как при использо-х таких процессов копии изготовляются на обычной бумаге, говечность этих копий весьма высока. Однако изображения остаются теплочувствительными. Поэтому при хранении ми. источника тепла изображения на них становятся

*Диазотипный процесс*, называемый также процессом ику и методом Геймса (по имени его изобретателя), основан ользовании селективного испарения жидкости с поверх-онируемого оригинала и ее конденсации на поверхности альных материала. Для копирования оригинала на его стороны наносится тонкая пленка легкоиспаряющегося атем на такой оригинал накладывается лист бумаги или материала и производится экспонирование ИК-лучами. копируе- мого изображения поглощают ИК-лучи и нагре-ется. Он- результате чего масляная пленка на них ких пары масла сразу же конденсируются на участках поверхности копировальной бумаги,

контактирующих с элементами изображения оригинала. На пробельных участках копируемого изображения такой перегонки (дистилляции) не происходит. Таким образом, на поверхности копировальной бумаги образуется скрытое изображение копируемого оригинала. Затем копировальная бумага отделяется от копируемого оригинала и подвергается проявлению. Проявление скрытого изображения производится путем напыления пигмента, который прилипает к тем участкам поверхности копировальной бумаги, где конденсировалось масло. Для фиксации проявленного позитивного изображения оно подвергается кратковременному ИК-облучению, приводящему к спеканию пигмента с бумагой. Копии получаются сухие, не боящиеся света, тепла и влаги.

Необходимо отметить, что при использовании дистилляционного процесса на поверхности копировальной бумаги образуется зеркальное изображение оригинала. Это означает, что для обеспечения возможности непосредственного чтения копии, снимаемой с непрозрачного оригинала, копию нужно изготовлять на прозрачной бумаге (например, кальке) и читать скопированный текст не с лицевой, а с обратной стороны. Такая копия на прозрачном материале может быть использована либо в качестве матрицы для размножения методом диазографии, либо для получения при помощи дистилляционного процесса копий с прямым изображением оригинала.

Дистилляционный процесс позволяет получать копии практически на любом материале — керамике, дереве, целлофане, резине, тканях, пластмассах и т. п. При изготовлении копии с непрозрачного оригинала на непрозрачном материале прямое изображение получается путем копирования первичной копии. Изображение на копии может быть любого цвета, даже из магнитного материала. Смачивание оригинала при копировании не оказывает на него никакого отрицательного воздействия. Дистилляционный процесс не требует ни затемненного помещения, ни дополнительной химической обработки копий.

Опытная копировальная установка, в которой применяется дистилляционный процесс, является ротационной и пока позволяет копировать лишь листовые оригиналы. Копирование производится на рулонную кальку, что делает ненужным преобразование копий с зеркальным изображением в копии с прямым изображением. Быстродействие установки — 21 *м/мин*, но может быть доведено до 61 *м/мин*.

### *Электрографические процессы*

Под электрографическими процессами принято понимать совокупность фотоэлектрических и электромагнитных способов формирования изображения и печати без натиска. К таким

процессам относятся электрография, электролитический процесс и термопластическая запись.

27.X.1961 г. русский изобретатель Е. Е. Горин впервые предложил использовать свойства фотополупроводников для воспроизведения изображения и подал заявку на изобретение под названием «Электрофотографический аппарат». Вероятно, Е. Е. Горин

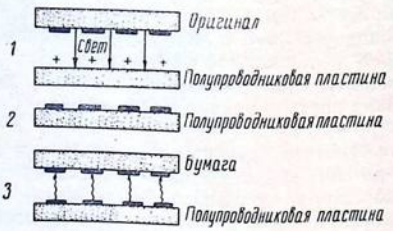



Рис. 110. Схема электрографического процесса с переносом изображения (ксерография)

*1* — отраженный от светлых участков оригинала свет деполяризует соответствующие участки заряженного положительным электричеством слоя фотополупроводника; *2* — отрицательно заряженные частицы пигмента притягиваются к участкам полупроводникового слоя, сохранившим положительный заряд; *3* — порошковое изображение переносится на бумагу и фиксируется на ней путем нагрева пигмента до температуры его плавления

впервые применил и термин «электрография». Американский физик Ч. Ф. Карлсон объединил принцип формирования изображения, предложенный Е. Е. Гориным, с принципом проявления скрытого электростатического изображения путем напыления красителя, предложенный профессором Будапештского университета П. Шеленьи, и получил 22.X.1938 г. первый электрографический оттиск.

Позже этот процесс был усовершенствован в лаборатории графики Мемориального института Баттеля (США) и назван ксерографией (от *греч.* «ксеро» — сухой, «графо» — пишу, рисую). Начало практического применения электрографии относится к 1950 г.

Электрография подразделяется на процессы с переносом и без переноса изображения. Схема *электрографического процесса с переносом изображения* приведена на рис. 110. На металлическую пластину нанесен слой фотополупроводника — селена, селенистого кадмия или сернистого кадмия, — имеющего удельное темновое сопротивление порядка $10^{15}$ *ом/см* и достаточно высокую фотоэлектрическую чувствительность. Сенсибилизация такой

пластины производится путем ее электризации при помощи коронного разряда. Напряжение на электродах разрядников обычно составляет 5—15 *кв*, а ток коронного разряда — 200—250 *мка*. Сенсибилизация пластины производится в темноте.

При экспонировании пластины с заряженным (электризованным) слоем полупроводника отдельные участки этого слоя деполяризуются пропорционально количеству света, воздействовавшего на данные участки. В результате этого на поверхности полупроводникового слоя образуется скрытое электростатическое изображение или потенциальный рельеф, соответствующий копируемому изображению. Экспонирование производится при помощи репродукционных аппаратов (в отраженном свете) или проекционных систем (в проходящем свете).

Для проявления скрытого электростатического изображения на поверхность полупроводникового слоя наносятся частицы красителя, которым предварительно сообщен заряд, противоположный по знаку заряду потенциального рельефа.

Частицы красителя притягиваются к заряженным участкам полупроводникового слоя и надежно удерживаются на его поверхности. В качестве красителей для электрографических процессов применяются измельченные синтетические или естественные смолы — копал, канифоль, сандарак, идитол, эфир гарпиуса, к которым зачастую прибавляются те или иные пигменты.

Наиболее простой способ электризации частичек красителя основан на так называемом трибоэлектрическом эффекте. Этот эффект состоит в электризации частичек двух разных диэлектриков зарядами противоположной полярности при трении этих частичек одна о другую. При этом знак и величина заряда определяются относительным расположением диэлектриков в трибоэлектрическом ряду. В литературе приводится следующий пример трибоэлектрического ряда: кукурузная мука, пробка, мел, пигментированный ликоподий, смола, песок, сернистая медь, винная кислота. При трении частичек этих диэлектриков одна о другую частички предшествующего в этом ряду диэлектрика получают положительный заряд, а частички последующего — отрицательный, причем величина заряда будет тем большей, чем дальше отстоят эти диэлектрики один от другого в данном ряду [2].

Трибоэлектрический эффект в электрографии используется следующим образом. Из двух компонентов — крупнозернистого носителя и мелкозернистого пигмента — составляется проявляющий порошок. В результате трения о носитель частицы пигмента приобретают соответствующий электрический заряд. Эти частицы пигмента оседают на поверхности частиц носителя, которые доставляют их заряженным участкам полупроводникового слоя.

[2] И. И. Жилевич и Е. Л. Немировский. Электрофотография. М., «Искусство», 1961, стр. 75.

В качестве носителя обычно применяют стеклянные или полисти... шарики размером 0,3—0,6 *мм*. Отношение количеств... к количеству носителя в проявляющем порошке состав... от 1 : 10 до 1 : 30.

Перенос проявленного изображения с полупроводникового слоя на обычную бумагу, ткань или какой-либо другой материал производится следующим образом. На порошковое изображение

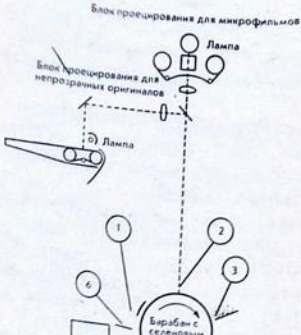



Рис. 111. Схема работы электрографического аппарата ротационного типа

*1* — электризация поверхности барабана; *2* — проецирование изображения на заряженную поверхность барабана; *3* — проявление скрытого изображения путем каскадной подачи проявителя; *4* — перенос отрицательно заряженного порошкового изображения с поверхности барабана на рулонную бумагу, заряжаемую положительным электричеством; *5* — закрепление изображения на бумаге расплавлением пигмента; *6* — очистка поверхности барабана

накладывается, например, лист бумаги, которому при помощи коронного разрядника сообщается электрический заряд, одноименный по знаку скрытого изображения и противоположный заряду порошкового изображения. В результате действия кулоновских сил частички пигмента притягиваются к поверхности бумаги и оседают на ней.

Фиксация порошкового изображения на поверхности бумаги производится либо путем нагревания пигмента до температуры его плавления, либо путем воздействия на пигмент парами растворителя — ацетона, четыреххлористого углерода, сернистого эфира и т. п. Возможна также фиксация порошкового изображения путем давления, для чего электрографический отпечаток пропускается между пресс-валиками.

Изготовление электрографической копии даже при ручном исполнении всех операций длится не более трех-четырех минут, а при использовании электрографических аппаратов ротационного типа — всего 10—15 сек. Процесс получения электрографического изображения осуществляется полностью сухим способом. Ис-

пользуемые в электрографии материалы дешевы и недефицитны. Для хранения материалов и осуществления процесса не нужно иметь затемненное помещение. Пластину с фотополупроводниковым покрытием можно использовать более 2000 раз, а изображения переносить не только на бумагу, но и на ткани, пленки и металлические формы для офсетной печати. Основным недостатком электрографии является плохое воспроизведение полутоновых изображений.

На рис. 111 приведена схема работы электрографического аппарата ротационного типа. В этом аппарате скрытое электростатическое изображение создается на поверхности непрерывно вращающегося барабана, покрытого слоем селена или другого фотополупроводника. На электризованный слой фотополупроводника динамически проецируется изображение копируемого оригинала. Проецирование производится либо с обычного микрофильма, либо отраженным светом (путем освещения непрозрачного оригинала). Проявление скрытого изображения осуществляется каскадным способом, который заключается в том, что двухкомпонентный проявляющий порошок ниспадает каскадом на поверхность барабана, покрытую слоем фотополупроводника и содержащую электростатическое изображение оригинала. Затем производится электростатический перенос порошкового изображения с поверхности барабана на рулонную или листовую бумагу и фиксация этого изображения на бумаге путем его нагрева до температуры плавления пигмента. Скорость работы электрографического аппарата ротационного типа при копировании на рулонную бумагу достигает 6 *м/мин*, что составляет примерно 25 копий нормального формата в 1 мин.

*Электрографический процесс без переноса изображения*, или метод «электрофакс» был разработан американской фирмой Radio Corporation of America в 1954 г. Этот процесс основан на использовании бумаги, покрытой слоем фотополупроводника (окиси цинка), на которой непосредственно изготовляются электрографические копии оригинала.

Электрографическая бумага изготовляется путем нанесения на подложку слоя диэлектрического связующего материала (силиконовой смолы, полистирола, сополимеров бутадиена и т. п.), в котором диспергирована окись цинка. Кроме того, в таком слое обычно содержатся органические красители, делающие окись цинка чувствительной к видимому свету (возможность такой сенсибилизации окиси цинка была установлена советскими исследователями академиком А. Н. Терениным и Е. К. Пуцейко).

Сенсибилизация электрографической бумаги и ее экспонирование производится таким же образом, как в электрографических процессах с переносом изображения. Электрографическая бумага заряжается отрицательным электричеством, а пигмент — положительным (рис. 112).

Проявление скрытого изображения на электрографической бумаге производится сухим или мокрым способом. Сухое проявление осуществляется методом «магнитной кисти». При использовании этого метода в качестве носителя в проявляющем порошке применяются железные опилки размером 0,25—0,5 *мм*. Эти опилки с осевшими на них частицами заряженного пигмента притягиваются электромагнитом, образуя некоторое подобие кисти. Если такой «кистью» провести по поверхности электрографической бумаги со скрытым изображением, то частицы пигмента притянутся к противоположно заряженным участкам

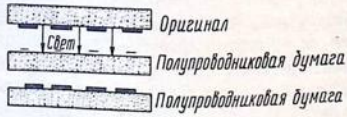



Рис. 112. Схема электрографического процесса без переноса изображения («электрофакс»)

*1* — отраженный от светлых участков оригинала свет деполяризует соответствующие участки заряженного отрицательным электричеством слоя фотополупроводника; *2* — положительно заряженные частицы пигмента притягиваются сохранившими отрицательный заряд участками фотополупроводникового слоя и закрепляются на них путем расплавления

фотополупроводникового слоя и проявят это скрытое изображение. Железные опилки останутся на поверхности якоря электромагнита. Методом «магнитной кисти» хорошо проявляются как штриховые, так и полутоновые изображения, а также изображения, имеющие большие сплошные участки. Фиксация порошкового изображения на электрографической бумаге производится путем нагрева пигмента до температуры его плавления.

Жидкое проявление осуществляется путем погружения экспонированной электрографической бумаги в проявитель или путем напыления последнего на поверхность такой бумаги. Проявитель представляет собой жидкость с высоким удельным сопротивлением (например, авиационный бензин, четыреххлористый углерод, фреон 113 и т. п.), в которой диспергирован соответствующий пигмент. В такой жидкости также происходит трибоэлектрическая электризация частиц пигмента. Оптимальная концентрация пигмента в жидкости составляет 1,5—2,5 *г/л*. Метод жидкого проявления обеспечивает получение высококачественных изображений с высокой оптической плотностью и с разрешающей способностью до 80 *л/мм* и более. Преимуществом жидкого метода проявления перед сухим является то, что он не требует последую-

щей фиксации отпечатков, а его недостатком — огнеопасность большинства используемых для проявления жидкостей.

Электрографические аппараты без переноса изображения обладают высоким быстродействием. Например, в сконструированном в НИИ электрографии (г. Вильнюс) аппарате ЭМА-1А бумажное полотно движется со скоростью 0,1 *м/мин*. В этом аппарате применяется жидкое проявление изображения.

Необходимо отметить, что если экспонирование электрографической бумаги производится в отраженном свете (копирование непрозрачного оригинала), то для получения на бумаге прямого (незеркального) изображения проекционно-оптическая система аппарата должна содержать зеркало.

Еще одной разновидностью электрографических процессов является *электролитический процесс*, который был разработан в 1957 г. американской фирмой Minnesota Minig and Manufacturing Co. Электролитический процесс предназначен для получения увеличенных позитивных копий с негативных 16- и 35-миллиметровых микрофильмов. Этот процесс основан на использовании специальной бумаги, на которую нанесена тонкая алюминиевая фольга, а на поверхность фольги — слой фотополупроводника, диспергированного в связующем веществе. В тех местах фотополупроводникового слоя, на которые попадает свет через пробельные участки негатива, соответствующих элементам изображения, удельное сопротивление фотополупроводника резко падает. Поэтому когда поверхность экспонированной электролитической бумаги смачивается электролитом — слабым раствором кислоты, содержащим серебро, а между фотополупроводниковым слоем и алюминиевой фольгой создается необходимая разность потенциалов, то на подвергнутых воздействию света участках фотополупроводникового слоя происходит процесс электролиза. В результате электролиза на белой поверхности такого слоя откладывается черное серебро, воспроизводя копируемое изображение.

Электролитический процесс применяется в читально-копировальных аппаратах типа «Фильмак-100, -200, -300 и -400». Для получения увеличенной позитивной копии с негативного микрофильма при помощи такого аппарата достаточно нажать кнопку. Продолжительность процесса копирования — 10 сек. Электролитическая бумага достаточно дорого стоит, а сохранность копий невелика. Но несмотря на это данный процесс получил широкое распространение.

Наконец, к числу электрографических процессов можно отнести также *термопластическую запись* графической информации [3]. Такая запись производится на пленке, которая состоит

[3] Термопластическая запись. Сборник переводных статей. М., «Искусство», 1966.

из подложки с высокой температурой плавления, покрытой прозрачным слоем электропроводящего материала. На этот электропроводящий слой нанесен слой термопластика (например, смесь дифенилсиликона с полимером полифениленоксида) толщиной от 5 до 15 *мк*, имеющего низкую температуру плавления. При облучении такой пленки модулированным электронным лучом на поверхности термопластика создается потенциальный рельеф, соответствующий структуре записываемого изображения.

Проявление скрытого изображения производится путем кратковременного нагрева термопластика до точки его плавления. Для этого в течение 0,05 сек. в электропроводящем слое пленки индуктивным путем создаются вихревые токи. Электростатические силы между зарядами на поверхности термопластика и на электропроводящем покрытии подложки вдавливают участки поверхности термопластика в тех местах, где имеются такие заряды. Вдавливание происходит до тех пор, пока силы электростатического притяжения не уравновесятся силами поверхност-

Таблица 32

**Стоимость и время копирования в различных процессах**
(по данным США)

| Процесс | Стоимость одной копии, *центы* а | Время изготовления одной копии, *сек* | Цена оборудования, долл. |
|---|---|---|---|
| Стабилизационный процесс . . . . | 5—8 | 10—15 | 100—2500 |
| Диффузионный процесс . . . . . | 6—9 | 30—60 | 50—650 |
| Матричный процесс («верифакс») . . | 10[в] | 40—60[в] | 132—549 |
| | 1[г] | 3—5[г] | |
| Диазографический процесс . . . . . | 1 | 5—25 | 130—11750 |
| Термографический процесс . . . . . | 4—5 | 5—6 | 220—595 |
| Электрографический процесс с переносом изображения (ксерография) | 5[д] | 30[в] | 29500[б] |
| | | 8[г] | |
| Электрографический процесс без переноса изображения («электрофакс») | 3—5[г] | 10—30[в] | 895—4475 |
| | | 4—10[г] | |

а) стоимость одной копии во всех процессах (за исключением электрографического с переносом и без переноса изображения) включает только стоимость бумаги и химикалиев; в двух последних процессах — включает также амортизацию оборудования или арендную плату за его использование;

б) арендная плата за использование машины: «Ксерокс-914» — 95 долл. в мес.; «Ксерокс-813» — 46 долл. в мес.;

в) первая копия;

г) каждая последующая копия;

д) при изготовлении не менее 200 копий в день.

ного натяжения термопластика. После этого пленка охлаждается ниже точки плавления термопластика, в результате чего деформации поверхности застывают и сохраняются в течение длительного времени. Запись производится электронной пушкой при вакууме около $10^{-4}$ *мм* рт. ст. Разрешающая способность термопластического слоя — 0,4 *мк*. Возможно последующее стирание записи.

Воспроизведение изображений, записанных на термопластической пленке, осуществляется в проходящем свете при помощи проекционной шлир-системы. В такой системе свет от источника, когда на поверхности пленки нет деформаций, фокусируется на экранирующие элементы многощелевой диафрагмы, вследствие чего экран остается неосвещенным. Когда же на пленке имеются деформации, то свет на этих деформациях рассеивается. Объектив проецирует рассеянные лучи света на экран, воспроизводя на нем записанное изображение.

Метод термопластической записи находится еще в стадии разработки, но возможные перспективы его применения в научно-информационной деятельности весьма широки.

* * *

При выборе наиболее подходящего процесса копирования, предназначенного для решения тех или иных задач, необходимо сначала установить, каким требованиям должен удовлетворять этот процесс. Для этого нужно попытаться ответить на следующие вопросы:

1. Каково основное назначение копировального аппарата.
2. Будет ли аппарат использоваться для выполнения каких-либо побочных работ. Если да, то насколько связаны основные и побочные работы. Не будет ли более целесообразным для выполнения вспомогательных работ иметь отдельный аппарат.
3. Каковы оригиналы, подлежащие копированию:
   а) гибкие или в переплете;
   б) одно- или двухсторонние;
   в) прозрачные или непрозрачные;
   г) цветные;
   д) сфальцованные или нет;
   е) какое изображение требуется, плотное или слабое.
4. Какое количество копий требуется:
   а) одновременно;
   б) в любой отрезок времени.
5. Каков максимальный размер копий.
6. Копируемые изображения штриховые или полутоновые.
7. Какова требуемая скорость получения копий в пиковый период.
8. Можно ли централизовать процесс копирования.
9. Имеется ли возможность использования типовых оригиналов.

Если такая возможность имеется, то:

   а) можно ли изготовлять матрицу в процессе подготовки самого документа;
   б) записываются ли одни и те же данные более одного раза;
   в) будут ли вноситься в копии какие-либо последующие дополнения, исправления или замены;

г) можно ли стандартизовать размеры копий, какие колебания допустимы в размерах копий;

д) нужно ли воспроизведение цветов оригинала; если да, то сколько требуется цветов.

10. Какое качество копии можно считать удовлетворительным.
11. Какова требуемая долговечность копии.
12. Может ли копия быть использована в качестве юридического документа.

Кроме факторов, относящихся непосредственно к копированию, следует учитывать также и факторы, относящиеся к размещению оборудования, и условия работы:

1. Какое помещение имеется для работы.
2. Требуется ли регулирование света и кондиционирование воздуха.
3. Нужна ли проточная вода.
4. Какова требуемая квалификация операторов.
5. Какие условия безопасности следует предусмотреть.

В зависимости от ответа на каждый из этих вопросов может оказаться целесообразным применение того или иного процесса копирования. Помимо этого, на выбор определенного процесса оказывают влияние экономичность и перспективность развития

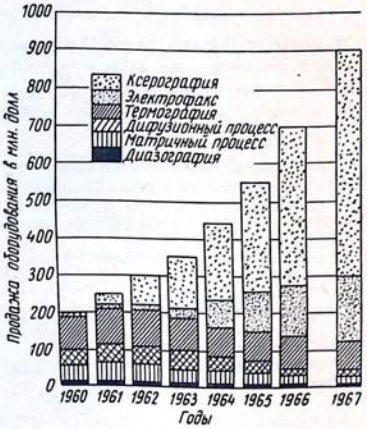



Рис. 113. Распространение оборудования для различных процессов копирования (по данным США)

тех или иных процессов. Естественно, что эти факторы в конкретных ситуациях могут зависеть от ряда побочных обстоятельств. Поэтому общих рекомендаций для выбора процесса копирования дать нельзя. Для облегчения этого выбора могут послужить дан-

ные [4] о сравнительных достоинствах и недостатках различных процессов копирования (табл. 31), о стоимости и времени копирования при использовании этих процессов (табл. 32), а также о распространении копировального оборудования (рис. 113).

## РАЗМНОЖЕНИЕ ДОКУМЕНТОВ

Размножением или тиражным размножением документов условно считается комплекс процессов изготовления свыше 15 копий со специальной печатной формы. Как уже было сказано, между процессами копирования и размножения нет четкой грани, так как некоторые процессы копирования, в первую очередь диазография, целесообразно использовать для получения тиража, превышающего 15 экземпляров, со специально подготовленного оригинала (печатной формы).

### *Типографские процессы*

Наиболее традиционным средством размножения документов являются типографские процессы, которые в отличие от более современных процессов размножения документов получили название классической полиграфии. Полиграфическое производство является самостоятельной отраслью знаний и потому не служит предметом рассмотрения в данной книге. Однако для обоснования процессов оперативной полиграфии необходимо конспективно наметить технологическую схему *издательских* и *типографских процессов* в классической полиграфии.

В общих чертах эти процессы сводятся к следующему (рис. 114). Рукопись любой работы, переданной автором в издательство, подвергается рецензированию для определения ее научной и практической ценности. Принятые к изданию рукописи тщательно редактируются с научной и литературной точек зрения. Редактор добивается ясного изложения основных мыслей автора работы, обеспечивает правильность терминологии и формулировок, лаконичность стиля, по возможности проверяет достоверность сообщаемых фактов и цифр. После редактирования рукопись перепечатывается на машинке. Затем производится вычитка рукописи корректором, отвечающим за правильность орфографии и пунктуации, и техническое редактирование, определяющее все полиграфические особенности будущего произведения печати: размер и рисунок шрифта, способ воспроизведения иллюстраций, формат набора и бумаги, качество обложки или переплета. На этом работа над рукописью в издательстве заканчивается, и она передается в типографию.

---

[4] A. Drattell. Making paper work. «Business Automation», 1967, v. 14, № 2, стр. 28—30; The revolution in office copying. «Chemical and Engineering News», 1964, v. 42, № 29, стр. 92.

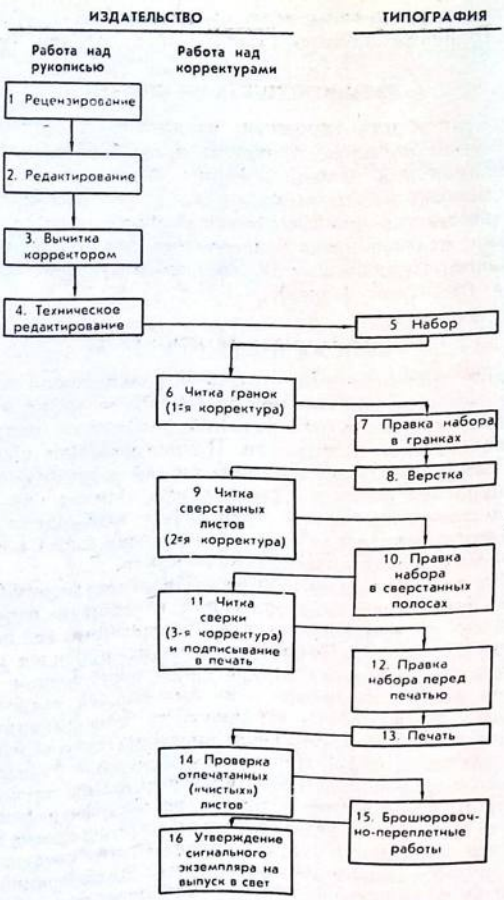



Рис. 114. Технологическая схема издательских и типографских процессов

Сложный полиграфический процесс в типографии можно разделить на три этапа: изготовление печатной формы (или набор), печатание и брошюровочно-переплетные работы. Различают три вида печати — высокую, глубокую и плоскую. При *высокой печати* печатающие элементы возвышаются над всей остальной поверхностью печатной формы. Этот вид печати применяется для создания основной массы книг, журналов, газет и других изданий, в которых преобладает текст. Печатная форма в этом случае состоит из набора, воспроизводящего текст, и клише, передающего иллюстрации.

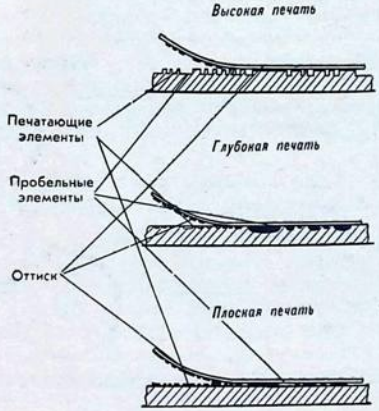
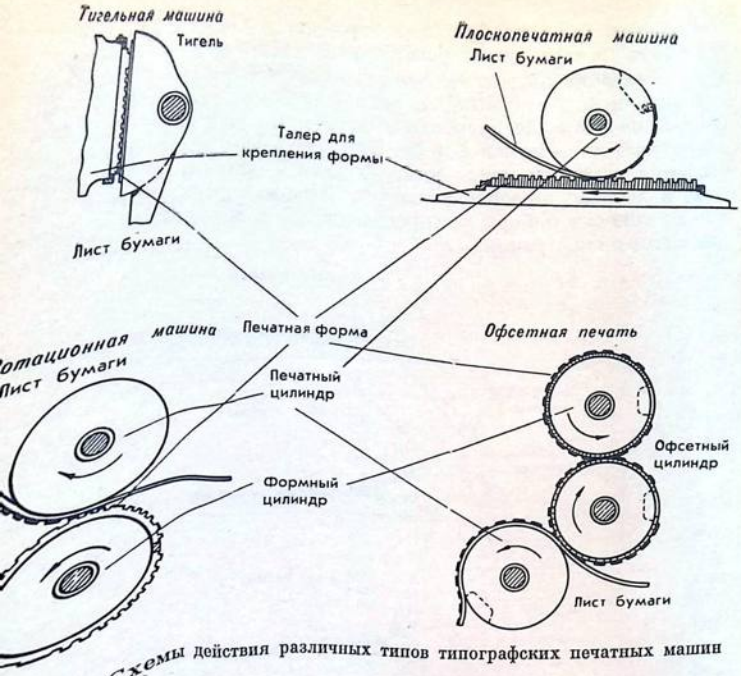



Рис. 115. Схемы различных видов печати

Типографский набор может осуществляться ручным способом, но чаще производится при помощи машин — строкоотливных (линотипов) или буквоотливных (монотипов). Строки набора соединяются в гранки, с которых делается предварительный оттиск, направляемый в издательство для прочтения. После исправления ошибок в гранках производится верстка — составление из набора и клише одинаковых по величине полос, содержащих материал страниц будущего произведения печати. Корректурные оттиски с верстки снова читаются в издательстве (а также обычно автором), а затем по ним делается правка набора. Этот процесс повторяется третий раз в так называемой сверке, после чего печатная форма считается готовой к печати.

В отличие от высокой печати при *глубокой печати* печатающие элементы углублены относительно пробельных элементов. Неодинаковая глубина печатающих элементов позволяет передавать




Схемы действия различных типов типографских печатных машин

116. ...менный переход от светлых мест оригинала к темным. Эта ...глубокой печати делает ее особенно ценной для передачи ...ественных иллюстраций. При *плоской печати* все печатающие ...непечатающие элементы лежат в одной плоскости. Этот ...ати — его наиболее известная разновидность называется ...афией — применяется обычно для многокрасочной печати ...плакатов, открыток и т. п. (рис. 115).

...особной формы производится печатание тиража, т. е. изготовление ...При первом способе перенос краски с печатной формы ...в момент непосредственного прижимания бумаги ...При втором способе (*офсетная печать*) краска с печатной ...нала переносится на промежуточный вал, а с него —

...временных типографиях печатание осуществляется при ...ечатных машин, которые подразделяются на тигельные, ...ные и ротационные. В *тигельных машинах* печатная

форма и поверхность, прижимающая бумагу (тигель), являются плоскостями. В *плоскопечатных машинах* плоскостью является только печатная форма, бумага к которой прижимается печатным цилиндром. В *ротационных машинах* и печатная форма, и поверхность, прижимающая бумагу, являются цилиндрами. Замена плоских поверхностей цилиндрическими позволяет намного ускорить работу печатных машин (рис. 116).

Работа над изданием в типографии заканчивается отделкой отпечатанных листов — их складывают в тетради (фальцуют), подбирают тетради в блок и сшивают; блок обрезают с трех сторон, покрывают обложкой или вставляют в переплет. Отделочные работы иначе называются *брошюровочно-переплетными*.

Совершенно ясно, что многочисленные операции по изданию произведения печати требуют значительного времени. Если учесть, что в издательстве и типографии работа ведется одновременно над многими изданиями, то становится понятным, что на выпуск каждого из них затрачивается по нескольку месяцев.

### *Подготовка печатной формы в оперативной полиграфии*

Потребность в средствах, которые позволяли бы размножать документальную информацию большими тиражами, но быстрее, дешевле и проще, чем традиционными средствами классической полиграфии, вызвала к жизни комплекс процессов, получивших название оперативной полиграфии. Этот комплекс состоит из комбинации средств безнаборной подготовки печатной формы, плоской офсетной, трафаретной или спиртовой печати, а также прогрессивной методики редакционно-издательской подготовки документов.

Появление оперативной полиграфии — принципиально новый шаг в развитии техники передачи информации, сравнимый по своему значению с изобретением книгопечатания. Достоинства оперативной полиграфии — экономичность, быстрота, портативность, доступность — позволяют при меньших затратах труда, материальных средств и времени достичь почти тех же количественных и качественных результатов, что и при классической полиграфии.

Преимущества оперативной полиграфии перед классической не сводятся только к сокращению производственного цикла, а следовательно, сроков издания за счет процессов корректуры. Авторская и издательская корректуры печатной формы являются необходимыми и важными этапами в любом полиграфическом процессе, так как существенно улучшают содержание издаваемых документов. Однако в оперативной полиграфии процессы корректуры осуществляются быстрее и проще благодаря применяемым в ней средствам подготовки печатной формы. Эти средства

Широко применяются наборно-пишущие машины «Флексорайтер» и «Джасторайтер» фирмы Frieden (США, Канада). На «Флексорайтере» одновременно изготавливают перфоленту и печатный текст, годный для получения офсетной формы. Обычный размер каретки 30,5 *см*. Шрифтовая клавиатура может быть одинарной или двойной. Управление «Флексорайтером» осуществляется автоматически при помощи перфоленты или карт с краевой перфорацией. «Флексорайтер» может быть установлен в комплексе с другими устройствами, например, с устройством для автоматического поиска данных на перфоленте, со счетными и адресовальными машинами, с электронно-вычислительными машинами, с системами дистанционной передачи информации.

Автоматическая наборно-пишущая машина «Джасторайтер» осуществляет автоматический набор текста с выключкой строк. «Джасторайтер» состоит из двух устройств — наборного автомата «рекордера» и воспроизводящего аппарата «репродуктора». Шрифт обоих устройств может быть разным, что увеличивает возможность шрифтовых выделений. Клавиатура имеет 43 клавиши, которыми при включении регистра можно набирать 86 знаков. На наборно-пишущем устройстве одновременно изготавливаются корректурный оттиск и перфорированная лента с соответствующим кодом для выключки строк. Перфолента устанавливается в воспроизводящий аппарат, который считывает ее и автоматически осуществляет набор с выключкой строки на обычной бумаге или непосредственно на бумажной офсетной форме. Скорость набора 100 слов в минуту. «Джасторайтер» позволяет также автоматически изготавливать «оригинал-макет»; на перфоленте можно закодировать не только текст, но и его расположение на «оригинале-макете». Таким образом, открывается возможность автоматизации двух основных этапов, предшествующих изготовлению офсетной формы, набора и изготовления «оригинала-макета». «Джасторайтер» сокращает стоимость набора на 80%. Стандартная длина каретки «рекордера» 30,5 *см*, максимальная длина строки около 23 *см*.

В СССР разработана и изготовлена автоматическая наборно-печатная машина «Север». Электронная машина «Север» предназначена для изготовления оригиналов, которые пополосно и построчно совпадают с набором, а также для автоматического кодирования и корректуры печатаемого текста с одновременным получением перфоленты для автоматического набора. «Север» состоит из трех устройств: наборно-программирующего, автоматического корректирующего и автоматического печатающего. Наборно-программирующее устройство включает электрическую наборно-пишущую машинку и электронный блок. Оно предназначено для одновременного набора текста и получения перфорированной ленты. Устройство работает на двух гарнитурах шрифта — академической и газетной, производит выключку стро-

ки на любой формат в пределах 1,5—7 квадратов. В случае, когда строка переполнена, устройство может вычитáть до правильного грамматического переноса. Производительность наборно-программирующего устройства до 120 тыс. печатных знаков, т. е. до 3 авторских листов, за 7-часовую смену. Автоматическое корректирующее устройство предназначено для многократного автоматического внесения правки и получения откорректированной перфоленты. Устройство включает записывающее и два читающих аппарата, электронный логический блок. Оно указывает пропущенные и выбирает лишние строки, подсчитывает количество строк. На автоматическом корректирующем устройстве после корректуры изготовляют ленту заборки. Автоматическое буквопечатающее устройство включает автоматическую пишущую машинку, трансмиттер, электронный блок. Оно обеспечивает получение готового текста с откорректированной перфоленты. Цикл работы автоматической системы «Север» заключается в следующем. Отредактированный издательский оригинал, перепечатанный на специальных форматных листах с занумерованными строками, набирается на наборно-программирующем устройстве. Одновременно с набором текста получают перфорированную ленту. После корректуры текста производится заборка исправленных строк; в случае необходимости делается заборка на заборку. Одновременно с набором исправленных строк получают перфорированную ленту. По основной перфоленте и перфоленте с текстом заборки изготавливают окончательно откорректированную ленту, которая вводится в линотип Н-10 для автоматического управления наборным процессом.

Наряду с наборно-пишущими машинами для изготовления печатных форм все шире начинают применяться фотонаборные машины. Фотонаборные машины имеют знаки различных алфавитов и гарнитур, включая математические, химические, астрономические символы; они обладают большими возможностями по изменению размеров шрифта. При наборе знаки экспонируются на фотопленку или фотобумагу, образуя негативное или позитивное изображение текста. Изображение текста после корректуры, правки и верстки в фотооригинал-макет копируется на печатную форму. В настоящее время широко применяются фотонаборные машины «Фотон» и «Линофильм» (США).

Машина «Фотон-200» состоит из электрической пишущей машины, фотосекции и устройства для управления специальными видами набора. Все части машины «Фотон» работают автоматически, за исключением пишущей машины, управляемой оператором. При нажатии на клавиши электрической машины посылаются электрические импульсы в специальное запоминающее устройство, которое последовательно суммирует ширину знаков и пробелов для автоматической выключки строк. После набора строки это устройство автоматически управляет работой фото-

657

секции, производящей экспонирование набранных знаков строки. В это время наборщик продолжает набор следующей строки, которая воспринимается вторым запоминающим устройством. При наборе следующих строк эти системы переключаются, т. е. роли их последовательно меняются. Изображение буквы проецируется через один из 12 объективов фотосекции машины. Экспонирование осуществляется при помощи мгновенной (4 *мксек*) вспышки, даваемой стробоскопической трубкой. На машине можно получить текст любой гарнитуры в 12 различных кеглях. При фотонаборе большие трудности представляет корректура и внесение исправлений на фотопленку. Исправленные строки заново набирают, проявляют, сушат и вставляют в полосу при помощи липкой ленты. Иногда перебирают целые страницы текста. Во избежание правки фотопленки на машине «Фотон» предусмотрена возможность корректуры набора до фотографирования. Одновременно с набором текста на электрической машине и фиксацией строки в запоминающем устройстве осуществляется набор строки на корректурном листе контрольной пишущей машины. После внесения необходимых исправлений строка фотографируется. На специальной панели производится верстка набранных полос. По макету верстки наборщик-оператор при наборе визуально проверяет расположение шрифта на фотопленке. Имеется устройство для пробивки набираемого и откорректированного текста на перфоленту, которая может быть использована для дистанционной передачи набора и автоматического дистанционного управления работой наборных машин.

Получают распространение малогабаритные устройства на основе наборно-пишущих и фотонаборных машин. Фирма «Америкен Тайп Фаундерс АТФ» (США) создала комплексы устройств «АТФ-Тайпсеттер» и «Альфатайп». В комплект «АТФ-Тайпсеттер» входит электрическая наборно-пишущая машинка «Джасторайтер» и буквопроецирующий аппарат, имеющий дополнительную клавиатуру, оптическую систему, наборный диск и регулировочные устройства. На машинке изготовляются перфолента и печатный текст. После корректуры перфолента вводится в буквопроецирующий аппарат, который автоматически производит фотонабор со скоростью до 200 знаков в минуту. В качестве носителя программы набора в машине «Альфатайп» применена, вместо перфорированной, магнитная лента, используемая многократно. Комплект «Альфатайп» состоит из электрической пишущей машинки, счетно-кодирующего блока и фотографирующего автомата, управляемого при помощи магнитной ленты.

Развитие средств подготовки печатной формы в оперативной полиграфии идет по пути создания электронных печатающих устройств. Быстродействующие печатающие устройства применяются для вывода текстовой и графической информации из ЭЦВМ.

Описанные выше средства позволяют получить лишь исходный оригинал для последующего изготовления печатной формы. Методы изготовления этой формы различны в зависимости от используемого в оперативной полиграфии вида печати. Наиболее часто применяют для размножения документов спиртовую, трафаретную и офсетную печать.

### *Спиртовая печать*

При изготовлении форм для спиртовой печати бумага или другой влагоустойчивый материал вводится в соприкосновение со специальной копировальной лентой или бумагой, покрытой щелочнорастворимым красителем. По мере того, как на оригинал наносится изображение, с обратной его стороны под действием красителя образуется обратное изображение. Сдавливающее усилие должно обеспечить прилипание красителя к обратной стороне бумаги. Процесс печати заключается в увлажнении бумаги спиртом, который растворяет краситель на обратной стороне оригинала и переносит некоторое количество краски на оттиск.

Аппарат для спиртовой печати называется гектографом. Он может полностью обслуживаться вручную, но имеются и аппараты, частично автоматизированные: с электроприводом и самонакладом. В гектографе увлажненный лист бумаги при помощи валиков прижимается к оригиналу с обратным изображением. Наилучшие оттиски получаются на писчей бумаге, которая обладает необходимой и достаточной впитывающей способностью. Применение других сортов бумаги ведет к тому, что изображение на оттиске получается слишком бледным или расплывчатым. Регулируя степень увлажнения бумаги и давления на нее в момент соприкосновения с печатной формой, можно изменять интенсивность изображения на оттиске и количество оттисков, получаемых с одной печатной формы. Это количество колеблется в пределах от 100 до 300 оттисков удовлетворительного качества. Спиртовая печать — единственный метод оперативной полиграфии, позволяющий получать многокрасочные оттиски за один прогон. Это достигается применением копировальной бумаги с красителями различных цветов. Однако наиболее устойчивое и интенсивное изображение дает краситель фиолетового цвета.

Оттиски, полученные при помощи спиртовой печати, быстро выцветают на свету и легко растворяются в воде. Поэтому спиртовую печать считают пачкающим процессом. Однако процесс этот сравнительно дешев, не требует специальной квалификации занятых в нем людей и сложного оборудования.

В настоящее время имеются гектографы со сменным штемпельным механизмом, которые позволяют получать оттиски спиртовой печати с форм, на которых меняются отдельные данные. Они используются для изготовления тиража бланков с разным

заполнением отдельных граф, например бланков заказа, учета, накладных счетов и т. п. В этих гектографах последовательно чередуются печатная форма с постоянным изображением и штемпельный механизм со сведениями, меняющимися для каждого оттиска.

### *Трафаретная печать*

Сущность метода трафаретной печати состоит в продавливании краски на поверхность бумаги через печатную форму — «трафарет». Процесс изготовления трафаретной формы заключается в формировании печатающих элементов в виде мельчайших отверстий, пропускающих краску через печатную форму, и в создании красконепроницаемых пробельных элементов.

Трафаретная форма применяется на ротаторах — простых и распространенных множительных аппаратах трафаретной печати. В качестве трафаретных форм используются «восковки», представляющие собой длинноволокнистую сетчатую ткань (шелковку), пропитанную парафином, воском или другим красконепроницаемым составом. Восковку устанавливают на каретку пишущей машинки (без красящей ленты). При ударе литерного молоточка разрушается красконепроницаемый слой и таким образом получается трафаретная форма, употребляемая для печати на ротаторе. При печати краска проходит через разрушения листа трафарета на воспринимающую поверхность — бумагу.

Благодаря простоте технологического цикла, экономичности применяемого оборудования и обслуживания трафаретная печать имеет большое будущее для оперативного размножения малотиражных изданий. Однако у этого вида печати имеются и недостатки.

Машинописные трафаретные формы на восковках не позволяют размножать иллюстративные материалы. Красконепроницаемый слой восковки разрушается даже при незначительном повышении температуры (свыше 20°) и слабом механическом воздействии; не отвечают современным требованиям и ротаторы, а также печатные краски для трафаретной печати и другие материалы.

Улучшение качества трафаретной печати, автоматизация процесса изготовления трафаретной формы, получение не только текстовых, но и иллюстрационных форм достигаются применением электроискрового способа и электронных устройств.

При электроискровом копировании на пленке, в местах изображения, угольное связующее вещество выгорает и оседает на гидрофильной поверхности печатной формы. В результате нагрева до 150—180° в течение 2—3 мин. образуется надежный по прочности рельеф. Тиражеустойчивость печатной формы, выполненной электроискровым способом, составляет несколько тысяч экземпляров.

Электронные копировальные аппараты отличаются высоким качеством воспроизведения оригиналов, оперативностью и экономичностью работы. Трафаретная печать с использованием таких аппаратов получила широкое распространение для размножения малотиражных информационных изданий. На электронно-копировальном аппарате можно изготовить трафарет с любого типографского, машинописного или рукописного оригинала. Изготовление трафарета происходит на двух барабанах, вращающихся с одинаковой скоростью. На одном из них находится оригинал, а на другом — трафарет. Фотоэлемент считывающего устройства воспринимает изображения текста, преобразуемые в токи высокой частоты и передаваемые на специальную токопроводящую ротопленку.

Ротаторы, как и гектографы, принадлежат к числу несложных в обращении и относительно недорогих аппаратов. Имеются ротаторы с электроприводом и самонакладом. Они позволяют изготовить до 4 тыс. оттисков в час. При этом одна печатная форма выдерживает несколько тысяч прогонов. Особенностью трафаретной печати является большой расход печатной краски и необходимость применения гигроскопической бумаги — типа почтовой или офсетной. Оттиски, получаемые при трафаретной печати, более долговечны и лучшего качества, чем получаемые при спиртовой печати, однако стоимость их несколько выше.

### *Офсетная печать*

Офсетная печать, как уже было сказано выше, — это способ, широко применяемый в классической полиграфии и основанный на передаче краски с печатной формы на бумагу через промежуточную эластичную поверхность (резиновый вал). Схема офсетной печати показана на рис. 116. Офсетная печать, обычно сочетающаяся с плоской, обладает рядом преимуществ перед печатанием с непосредственным контактом бумаги с печатной формой. Применение этого способа уменьшает износ печатной формы и повышает скорость печатания. В оперативной полиграфии использование офсетной печати обусловлено возможностью быстрого получения относительно больших тиражей (в несколько тысяч экземпляров) с отпечатками высокого качества при сравнительной простоте изготовления печатной формы.

Для изготовления офсетных форм применяется дефицитная алюминиевая фольга, которая используется три-четыре раза. При повторном использовании фольга подвергается зернению, сложным химическим процессам. Когда на фольгу непосредственно печатают текст, она используется только один раз. Такое использование фольги экономически невыгодно. Для печати небольших тиражей информационных изданий применяют печатные формы на гидрофильной бумаге. На изготовление таких форм расходуется

меньше средств, а процесс изготовления ускоряется более чем в два раза. Печатные формы на гидрофильной бумаге легко сохранить для допечатки тиража; они могут изготовляться на пишущей машине. Прокаливание бумаги инфракрасными лучами позволяет увеличить тиражеустойчивость формы с 150 до 2500 экземпляров.

Особенно перспективны электрографические методы изготовления офсетных форм. В одном случае на промежуточных фотополупроводниковых слоях получают изображение, которое электрическим способом переносится на металлическую или гидрофильнобумажную офсетную форму. В другом случае изображение формируется без промежуточных материалов, непосредственно на поверхности офсетной формы, покрытой полупроводниковым и гидрофильным слоем.

Для офсетной печати в оперативной полиграфии применяют упрощенные малоформатные печатные машины, называемые ротапринтами. Существуют модели ротапринтов, рассчитанные на формат писчего листа, половины писчего листа и двойного писчего листа. Ротапринты позволяют размножать материалы текстового и иллюстративного характера (штриховые и полутоновые иллюстрации) со скоростью около 5 тыс. односторонних однокрасочных оттисков в час. Многокрасочная печать осуществляется путем последовательных однокрасочных прогонов.

Этот способ размножения документов может быть отнесен к числу наиболее дорогих и сложных. Обслуживание ротапринта и подготовка печатных форм для него требует специальной квалификации. Стоимость ротапринта превышает стоимость оборудования для других способов размножения. Офсетная печать требует специальной офсетной бумаги, обладающей особыми качествами. Такая бумага должна быть влагоустойчивой, прочной, отличаться белизной и малой пылимостью. В связи с двукратной передачей краски при офсетном способе печати толщина слоя краски на оттиске меньше обычного, что влечет необходимость применения красок повышенной интенсивности.

Вместе с тем это наиболее совершенный из всех способов размножения документов в оперативной полиграфии. Он широко используется в тех случаях, когда имеется постоянная потребность в получении большого числа копий высокого качества, для выпуска вторичных информационных изданий: экспресс-информации, сигнальной информации и т. п.

### *Размножение на адресовальных машинах*

Адресовальные машины (адресографы, адремы) применяются главным образом для печатания фамилий и адресов подписчиков на периодических изданиях. Они позволяют многократно воспроизводить в установленной последовательности определенный

Таблица 33

**Выбор средств размножения документов ***

| Критерии выбора / Применяемые средства | | Пишущая машина | Автоматическая пишущая машина | Гектограф (спиртовая печать) | Гектограф со сменным штемпельным механизмом | Ротатор с машинописным трафаретом («восковкой») | Ротатор с электроискровым трафаретом | Ротапринт (офсетная печать) | Адресовальная машина |
|---|---|---|---|---|---|---|---|---|---|
| Оттиск | Удовлетворительного качества | | | X | X | | | | |
| | Хорошего качества | X | | | | X | | | X |
| | Высокого качества | | X | | | | X | X | |
| Текст | Машинописный | X | X | | X | X | | | X |
| | Любого вида | | | X | / | | X | X | |
| Иллюстрации | Штриховые | | | X | | / | X | X | |
| | Тонкие штриховые (машиностроительные чертежи) | | | | | | X | X | |
| | Полутоновые | | | | | | X | X | |

Таблица 33 (*окончание*)

| Критерии выбора | | Применяемые средства: Пишущая машина | Автоматическая пишущая машина | Гектограф (спиртовая печать) | Гектограф со сменным штемпельным механизмом | Ротатор с машинописным трафаретом («восковкой») | Ротатор с электроискровым трафаретом | Ротапринт (офсетная печать) | Адресовальная машина |
|---|---|---|---|---|---|---|---|---|---|
| Цвет | Однокрасочная печать | X | X | X | X | X | X | X | X |
| | Многокрасочная печать (за один прогон) | | | X | / | | | | |
| | Многокрасочная печать (за несколько прогонов) | | | | | / | X | X | |
| Тираж | До 6 экз. | X | X | | | | | | X |
| | От 6 до 12 экз. | / | X | / | / | | | | X |
| | От 12 до 100 экз. | | X | X | X | | | | X |
| | От 100 до 300 экз. | | / | X | X | / | / | / | X |
| | От 300 до 3000 экз. | | | | | X | X | X | X |
| Бумага | Писчая | / | / | X | X | | | / | X |
| | Типографская | X | X | | | / | | | X |
| | Офсетная | | | | | / | / | X | X |
| | Почтовая | | | | | X | X | | X |

*) Обозначения: X — рекомендуется, / — допустимо

набор сведений. Однако эти машины могут использоваться и для размножения некоторых информационных материалов.

Имеются адресовальные машины, основанные на принципах спиртовой и трафаретной печати. Но чаще всего в этих машинах используется печать тиснением (разновидность безнаборной высокой печати). Печатная форма для адресографов изготовляется выдавливанием букв на пластмассовой, картонной или металлической пластине. Полученное таким образом рельефное обратное изображение текста через красящую ленту оттискивается на бумаге.

Достоинства размножения на адресовальных машинах способом тиснения заключаются в высоком качестве оттиска, возможности использования любых сортов бумаги, практически неограниченном тираже. Однако основное назначение этих машин обусловливает и их недостатки: малую величину печатной формы, умещающей лишь несколько строк текста, невозможность размножения иллюстративного материала, небольшую скорость печатания, трудоемкость изготовления печатной формы.

Адресовальные машины часто применяются в зарубежных библиотеках для размножения каталожных и библиографических карточек.

* * *

В данном разделе кратко рассмотрены лишь самые основные средства размножения информационных документов. Выбор того или иного из этих средств определяется многими критериями: необходимым качеством оттиска, характером оригинала (текст, иллюстрации), величиной тиража и т. п. При этом приходится учитывать объем множительных работ данного учреждения, возможность приобретения того или иного оборудования, квалификацию сотрудников, занятых множительными работами, и т. д. Поэтому приведенные выше указания на достоинства или недостатки различных способов носят самый общий характер. Некоторую помощь в выборе средств размножения может оказать табл. 33, составленная на основе опыта множительных работ в ГДР [5].

---

[5] H. Schöppenthau. Rationelles Vervielfaltigen. «Rechentechnik», 1966, Jg. 3, № 6, стр. 27.

Глава 9

# ОРГАНИЗАЦИЯ НАУЧНО-ИНФОРМАЦИОННОЙ ДЕЯТЕЛЬНОСТИ

## НАУЧНО-ИНФОРМАЦИОННАЯ ДЕЯТЕЛЬНОСТЬ В СССР

Великая Октябрьская социалистическая революция коренным образом изменила общественное положение и роль науки в нашей стране. С первых дней установления Советской власти В. И. Ленин и Коммунистическая партия уделяли большое внимание всем мероприятиям, которые обеспечивали развитие советской науки, использование новейших научно-технических достижений в интересах социалистического строительства.

В. И. Ленин указывал, что выполнение плана развития науки, экономики и образования, намеченного Программой партии, в значительной мере будет зависеть от творческого изучения и использования всего ценного, что имеется в научно-технических достижениях и производственном опыте передовых капиталистических стран. Он учил оценивать результаты теоретических и экспериментальных исследований советских научных учреждений и институтов с позиций мирового уровня достижений науки и техники. Поэтому, подчеркивал В. И. Ленин, нам надо «перенять все действительно ценное из европейской и американской науки; — это наша первейшая и главнейшая задача» [1].

Из общей совокупности мероприятий, направленных на обеспечение выполнения программных задач развития советской науки и создания на ее основе социалистической экономики, В. И. Ленин особо выделял задачу всестороннего изучения мировой литературы по всем отраслям знания с целью распространения новейших достижений науки и техники и внедрения их в социалистическое производство.

14. VI 1921 г. по инициативе В. И. Ленина был принят декрет Совета Народных Комиссаров «О порядке приобретения и распределения заграничной литературы». Этим декретом была учреждена

[1] В. И. Ленин. Полное собрание сочинений, т. 45, стр. 206.

специальная «Центральная междуведомственная комиссия по закупке и распределению заграничной литературы (Коминолит)».

В декрете указывалось, что задача Коминолита — получение из-за границы всякого рода литературы, необходимой для РСФСР по всем отраслям знания, в первую очередь вышедшей со второй половины 1914 г., сосредоточение всей заграничной литературы в соответствующих научных учреждениях и библиотеках, а также распределение и организация рационального использования заграничной литературы всеми учреждениями и отдельными лицами.

«О наиболее ценных политических, научных, научно-технических изданиях, — говорилось в декрете, — Коминолит публикует систематические указатели, с точным обозначением библиотек и хранилищ, где таковые издания находятся» [2].

Так Ленинским декретом от 14.VI 1921 г. было положено начало организации советской государственной системы научной информации. Коминолит был создан как первый научно-информационный орган Советского государства.

В. И. Ленин пристально следил за работой Коминолита, определял направление его деятельности. «Главная задача, — писал он, — которую должен себе поставить Коминолит, — это добиться того, что в Москве, Петрограде и крупных городах Республики было сосредоточено в специальных библиотеках по 1 экземпляру всех заграничных новейших технических и научных (химия, физика, электротехника, медицина, статистика, экономика и пр.) журналов и книг 1914—1921 гг. и было бы налажено регулярное получение всех периодических изданий. Всю работу Коминолита я буду оценивать в первую очередь с точки зрения реального выполнения этого задания» [3].

В. И. Ленин придавал научной информации настолько важное значение, что считал наблюдение за переводом и изданием мировой научной литературы одной из основных функций заместителей председателя Совета Народных Комиссаров и Совета Труда и Обороны. Так, в постановлении о работе замов (апрель 1922 г.), подготовленном В. И. Лениным, указывается: «. . . под надзором замов должна быть переведена и издана вся лучшая новейшая литература, особенно американская и немецкая, об организации труда и управления» [4].

Не случайно поэтому и в статье «О значении воинствующего материализма», опубликованной в марте 1922 г. в ежемесячном философском журнале «Под знаменем марксизма», В. И. Ленин вновь обращает внимание на важность осуществления перевода и реферирования мировой научной литературы, а также на большое значение публикации обзорной информации.

[2] «Собрание узаконений и распоряжений Рабочего и Крестьянского Правительства», 1921, № 51, стр. 285.
[3] В. И. Ленин. Полное собрание сочинений, т. 53, стр. 228—229.
[4] Там же, т. 45, стр. 154—155.

сам библиотековедения и библиографии в 1936 г. эти принципы подверглись всестороннему обсуждению.

В 1936 г. по специальному указанию ЦК партии была осуществлена дальнейшая централизация библиографической работы. Центром текущей информации по технике стала Государственная научная библиотека, приступившая к изданию ежемесячного журнала «Новости технической литературы» («НТЛ») (1936—1953 гг.). Журнал включал шесть серий, охватывавших горную промышленность, металлургию и технологию металлов, машиностроение, строительную промышленность, химию и химическую промышленность, энергетику и энергопромышленность. Реферированию и библиографированию подвергалась вся советская техническая периодика, около 1,5 тыс. названий зарубежных технических журналов, технические каталоги, неопубликованные труды научно-исследовательских институтов и другие материалы. Поскольку за пределами «НТЛ» остались легкая промышленность, транспорт, механизация сельского хозяйства, информационную работу по этим отраслям науки Академия наук СССР планировала организовать В области несколько серий реферативного журнала, из которых выпуск вышли два — «Физико-математический реферативный журнал» (1938—1941 гг.) и «Химический реферативный журнал» (1938—1941 гг.).

Ценный материал по истории научно-информационной деятельности в нашей стране содержится в изданной ВИНИТИ книге Ю. Теплова [7], из которой почерпнуты приведенные выше сведения.

Во все годы создания и развития Советского государства Коммунистическая партия уделяла большое внимание делу организации научно-информационной деятельности в нашей стране. Но наиболее интенсивное развитие научно-информационная деятельность получает у нас в послевоенные годы (1945—1950 гг.).

К 1950 гг. информационное обслуживание научных работников и специалистов во всем народном хозяйстве СССР обеспечивается главным образом через сеть научно-технических библиотек. Наращивание темпов прогресса науки и техники заставило организационные формы научно-информационной деятельности, так как научно-технические библиотеки с их традиционными методами работы перестали удовлетворять быстро возросшие потребности в научно-технической информации при непрерывном росте количества информационных источников.

Новые требования привели к созданию в стране разветвленной системы научно-технической информации. Такие органы

---

Ю. Теплов. Информационно-техническая библиография в СССР. Исторический очерк. М.—Л., 1962. 124 стр. с илл. (Труды ВИНИТИ).

появились в отдельных отраслях промышленности, в научных учреждениях, на промышленных предприятиях.

В 1952 г. по инициативе академика А. Н. Несмеянова Президиум АН СССР внес предложение в Правительство СССР об организации Центра научной информации. В том же году в системе Академии наук был создан Институт научной информации, который в последующие годы вырос в головной центр научно-информационной деятельности в нашей стране. Основные функции и содержание его работы будут изложены ниже.

Как показала практика, отдельные изолированные информационные ячейки не могут обеспечить эффективной научно-информационной деятельности и прежде всего — исчерпывающей полноты информации по мировым источникам.

В стремлении обеспечить своих потребителей как можно бо́льшим количеством сведений информационные органы допускали в больших объемах ничем не оправданное дублирование. Строго говоря, в стране еще не было единой государственной информационной службы, четкого разграничения целей, задач и функций отдельных информационных звеньев. Это требовало перестройки всей научно-информационной деятельности в стране и прежде всего строгой государственной координации этой деятельности.

Предусмотренное Программой КПСС всемерное содействие дальнейшему усилению роли науки в строительстве коммунистического общества, образцовой постановки научно-технической информации, всей системы изучения и распространения отечественного и зарубежного передового опыта в последние годы нашло свое отражение в ряде последовательных Постановлений ЦК КПСС и Совета Министров СССР, направленных на создание у нас в стране образцовой системы научно-технической информации.

Начиная с 11.V 1962 г., последовательно, шаг за шагом, улучшалась информационная деятельность в стране. Она приобретала более четкие организационные формы, способствующие повышению ее эффективности.

На XXIII съезде КПСС тов. А. Н. Косыгин говорил: «Технический прогресс в народном хозяйстве и успехи науки в большой степени зависят от хорошо поставленной системы информации о результатах научных исследований, проводимых в стране и за рубежом, о достижениях и новых методах производства, об изобретениях и рационализаторских предложениях.

Нам необходимо создать в стране высокоэффективную общегосударственную систему научной информации. Она поможет плановым, хозяйственным органам и научно-исследовательским организациям ускорить внедрение в народное хозяйство научно-технических достижений. Наши исследователи и конструкторы

ленного проникновения в сущность изучаемого предмета («ученый как бы копает колодец»). С другой стороны, происходит синтез научных знаний, основанный на сопоставлении результатов, научных фактов, полученных в различных, казалось бы не связанных между собой отраслях знаний. На стыках между отраслями науки возникают новые научные направления, происходит взаимопроникновение наук, их предметов и методов исследований. Эти тенденции влияют на систему научных публикаций таким образом, что литература по одному вопросу или отрасли знания оказывается рассеянной в очень широком круге изданий.

Условия централизованной обработки научной информации позволяют с наименьшими затратами достичь наиболее полного охвата мировых источников информации по различным направлениям исследований, систематизировать и обобщить сведения, разбросанные в огромном числе этих источников. Это положение можно проиллюстрировать, например, на двух сериях советского Реферативного журнала (РЖ) «Химия» и «Биология». Число названий профильных журналов, относящихся по своему содержанию полностью к области химии и химической технологии, достигает на всех языках во всех странах мира примерно 700—750. В то же время в РЖ «Химия» освещаются публикации приблизительно из 12—13 тыс. названий периодических изданий. То же самое можно видеть и в серии РЖ «Биология», где число профильных журналов составляет примерно 800 названий, а число периодических изданий, используемых для реферирования, составляет 5,5—6 тыс. названий.

Американские специалисты в области научной информации фактически стоят на противоположных позициях относительно принципов построения системы научно-технической информации, хотя и не отрицают целесообразности создания крупных научно-информационных органов в каждой отрасли науки и техники. В качестве таких информационных органов, обслуживающих отдельные отрасли, можно назвать институты Chemical Abstracts Service, Biological Abstracts, Inc., Национальную медицинскую библиотеку и т. д.

Вместе с тем в последние годы американские специалисты все чаще высказывают мысль о необходимости координации государством деятельности отдельных разрозненных информационных органов. Нам хотелось бы привести мнение самих американских специалистов по этому вопросу. «Вызывает сомнение предложение о том, что государству следует взять на себя сбор, хранение и распространение информации, дополняемое утверждением, будто с помощью информационно-вычислительных машин можно достигнуть огромного успеха. В этом плане представляет интерес позиция комиссии сената. Комиссия сената исходит из того, что не следует формировать централизованную

систему информации, а нужно создать нечто вроде палаты связи или расчетной палаты различных систем или узлов взаимодействующих систем информации. Масса научной и технической информации слишком велика, чтобы ее можно было централизовать в каком-либо месте. Сейчас существует слишком много раздельно действующих систем и нет никакой возможности оторвать существующие системы информации от органов, создавших эти системы. Это делать и нежелательно. Но надо надеяться, что все многообразие существующих систем будет связано, что между ними установятся также связи и отношения, когда любой исследователь, в каком бы пункте страны он ни находился, сможет включиться в эту систему так, как он включается в телефонную сеть. Будет создан единый центр, откуда можно будет получить направление в соответствующее хранилище информации» [9].

Разделение функций между органами научно-технической информации в СССР основывается на следующих принципах. Все научно-информационные источники, т. е. литература периодического и разового издания (журналы, книги, отчеты о научных совещаниях и научных исследованиях, патенты) и все сведения (технологического, технического содержания) о практике в промышленности, в сельском хозяйстве, в области транспорта и т. д. могут быть разделены на два самостоятельных информационных потока (рис. 117).

Поток информации, получаемой путем централизованной обработки мировой научно-технической литературы, «течет» сверху вниз, питая все ячейки информационной службы страны, а также отдельных потребителей. Этот поток получил название нисходящего.

Второй поток образуют непубликуемые документы, в которых отражаются результаты научно-исследовательских и опытно-конструкторских работ, технологический опыт, рационализаторские предложения и т. д. Такие документы, как правило, носят служебный характер и поступают из низовых ячеек в центральные отраслевые органы, где они синтезируются и получают соответствующее освещение в изданиях информационного органа данной отрасли. Этот поток получил название восходящего.

Сеть информационных органов СССР включает (рис. 118):

1. Всесоюзные информационные институты, к числу которых относятся:

Всесоюзный институт научной и технической информации Государственного комитета по науке и технике СССР и Академии наук СССР (ВИНИТИ);

---

[9] Наука — техника — управление. Интеграция науки, техники и технологии, организации и управления в Соединенных Штатах Америки. Под ред. Ф. Каста и Д. Розенцвейга. М., «Сов. радио», 1966, стр. 449—450.

На отраслевые министерства и ведомства возложена ответственность за развитие научно-информационной деятельности в отрасли, а на Советы Министров союзных республик — за развитие органов научно-технической информации в республике, а также на предприятиях и в организациях республиканского подчинения.

Это Постановление комплексно решило вопрос о работе всей сети органов научно-технической информации и пропаганды как единого механизма, работающего по строго координируемому плану. Устранена разобщенность в работе всесоюзных, центральных, отраслевых, республиканских институтов информации, отделов и бюро информации предприятий, научно-исследовательских, проектно-конструкторских организаций, научно-технических библиотек, технических кабинетов. Четко разграничены функциональные обязанности между всеми органами научно-технической информации и пропаганды. Установлена взаимосвязь в работе органов информации и научно-технических библиотек по созданию справочно-информационных фондов, внедрению систем классификации научно-технической литературы и информационных материалов, комплектованию фондов отечественной и иностранной литературы, методическому руководству сетью библиотек, комплексной механизации и автоматизации подготовки и поиска информации.

В связи с усилением деятельности отраслевых органов информации определен порядок подготовки и представления предприятиями и организациями научно-технической, производственной и экономической информации в центральные отраслевые информационные органы и выдачи последними различных информационных материалов по требованиям заинтересованных организаций. Установлены виды и тематика изданий органов научно-технической информации, в значительной мере исключающие дублирование в их работе. Определена роль домов научно-технической пропаганды, домов техники и технических кабинетов в общей системе мероприятий по научно-технической информации и пропаганде передового производственно-технического опыта.

## *Всесоюзные органы научно-технической информации*

### Всесоюзный институт научной и технической информации (ВИНИТИ)

19. VII 1952 г. Совет Министров СССР принял постановление об организации Института научной информации в системе Академии наук СССР. На Институт были возложены следующие задачи: систематическое и исчерпывающее реферирование всей мировой литературы в области естествознания и техники (астро-

номии, геодезии, математики, механики, кибернетики, физики, химии, биохимии, биологии, геофизики, географии, геологии, горного дела, машиностроения, транспорта, автоматики, радиоэлектроники, электротехники, энергетики, экономики промышленности), подготовка и издание на этой основе реферативного журнала; выпуск обзорно-библиографической и справочной литературы, экспресс-информации по наиболее актуальным темам, а также организация и развитие научных исследований, направленных на совершенствование существующих методов и технических средств, используемых в научно-информационной деятельности.

Для работы в создаваемом институте научной информации Академия наук СССР направила большое число крупных ученых и специалистов в различных областях науки и техники. Принимая во внимание большое значение, которое Институт научной информации приобрел для самых широких кругов научных и инженерно-технических работников, а также его ведущую роль в развитии научно-технического прогресса в стране, Совет Министров СССР в декабре 1955 г. преобразовал Институт научной информации во Всесоюзный институт научной и технической информации (ВИНИТИ), подчинив его одновременно Государственному комитету по науке и технике СССР и Академии наук СССР.

За полтора десятилетия существования ВИНИТИ прошел большой путь становления и организации, путь, связанный с трудностями накопления опыта, а главным образом — создания квалифицированного коллектива специалистов, насчитывающего сейчас более 2,5 тыс. штатных сотрудников (без учета 1,7 тыс. человек, занятых в издательском секторе и полиграфии). Кроме того, в настоящее время в ВИНИТИ работает более 22 тыс. внештатных научных сотрудников самых различных специальностей.

ВИНИТИ ведет работу в трех основных направлениях: научно-информационном, научно-исследовательском и научно-организационном. Выполняя поставленную перед ним задачу обеспечения специалистов исчерпывающей научной и технической информацией, ВИНИТИ выпускает Реферативный журнал, Экспресс-информацию, обзоры «Итоги науки» и другие информационные издания.

Реферативный журнал — наиболее фундаментальное информационное издание, предназначенное для текущей и ретроспективной информации. С 1960 г. Институт издает 16 серий единого Реферативного журнала Советского Союза.

Реферативный журнал ВИНИТИ охватывает естественные и технические науки, а также экономику промышленности и выходит в следующих сериях: «Автоматика и радиоэлектроника», «Астрономия и геодезия», «Биология», «Биологическая химия», «География», «Геология», «Горное дело», «Геофизика», «Матема-

тика», «Машиностроение», «Металлургия», «Механика», «Физика», «Химия», «Экономика промышленности» и «Электротехника».

Учитывая растущий интерес специалистов к вопросам теории, методики и практики научно-информационной деятельности, Институт с 1963 г. приступил к изданию отдельного выпуска РЖ «Научная и техническая информация», освещающего следующие вопросы: общая проблематика научной информации; вопросы информационного анализа, научный перевод (в том числе и машинный); информационно-поисковые системы; документальные источники информации, их виды, сбор, хранение и распространение; методы и средства копирования документов; оперативная полиграфия; общетехнические средства обеспечения информационной работы; использование информационных документов; организация научно-информационной деятельности. В соответствии с просьбой Международной Федерации документации (ФИД) в ответ на многочисленные пожелания зарубежных специалистов ВИНИТИ приступил с 1967 г. к выпуску серии РЖ «Научная и техническая информация» на английском языке.

Ценность Реферативного журнала ВИНИТИ определяется прежде всего полнотой охвата литературы по соответствующим отраслям науки и техники, т. е. огромным количеством источников информации, которые используются для его подготовки.

Все периодические и продолжающиеся издания — журналы; сборники; научные труды; ученые записки; вестники высших учебных заведений, научно-исследовательских организаций и национальных академий наук; отчеты международных конгрессов, конференций и симпозиумов; патентные описания; стандарты; книги; монографии; справочники; словари; энциклопедии и другие публикации, представляющие научную ценность, — приобретаются ВИНИТИ путем подписки, международного обмена или бесплатно (например, все отечественные издания). Языковое разнообразие научно-технической литературы, поступающей в ВИНИТИ, достаточно велико. В Институт поступает литература более чем на 70 языках (в том числе на 19 языках народов СССР). Первоисточники в ВИНИТИ поступают более чем из 100 стран мира.

Широкий охват документальных источников научной информации стал возможен благодаря систематическому изучению состояния и развития отечественной и зарубежной научно-технической литературы. Новые необходимые для ВИНИТИ публикации выявляются путем повседневного и тщательного изучения отечественных и зарубежных реферативных и библиографических изданий, проспектов, каталогов издательств и книготоргующих организаций, а также предложений различных иностранных научных организаций и высших учебных заведений, заинтересованных в обмене научными публикациями с ВИНИТИ.

В результате тщательного изучения всевозможных видов изданий удается следить за появлением новых источников информации. Практика последних лет показала, что в мире ежегодно возникает до 2,0—2,5 тыс. новых научно-технических журналов. Эти журналы вначале поступают в ВИНИТИ в виде образцов, по которым определяются их профиль и научная ценность. При этом отбор производится весьма строго. Так, в 1965 г. из 2967 просмотренных новых журналов были признаны ценными для реферирования лишь 787. Опыт ВИНИТИ по изучению мировой периодической литературы по естественным и техническим наукам в настоящее время обобщается в аннотированном указателе. Первый том этого указателя, посвященный периодическим и продолжающимся изданиям по астрономии, математике и механике, выйдет из печати в 1968 г., а последний, седьмой том, посвященный изданиям общего характера, — в 1969—1970 гг. В указатель будет включено более 15 тыс. наименований зарубежных периодических и продолжающихся изданий, реферируемых во всех сериях Реферативного журнала.

Непрерывный рост числа публикаций обусловливает увеличение объема Реферативного журнала: в 1967 г. объем всех серий Реферативного журнала составил уже 20 613 авторских листов. С целью повышения доступности Реферативного журнала для индивидуального читателя и обеспечения большего удобства работы с этим журналом Институт с 1958 г., наряду со сводными томами, начал издавать Реферативный журнал в виде отдельных выпусков по более узким отраслям, которые составляют разделы сводных томов отдельных серий журнала.

Полнота использования материалов, помещенных в Реферативном журнале, в значительной мере зависит от его оснащения справочным аппаратом. Поэтому Реферативный журнал снабжается номерными и годовыми авторскими указателями, а также годовыми предметными указателями и, кроме того, для РЖ «Химия» — систематическими, формульными и патентными указателями.

В качестве справочного аппарата широко используется система ссылок внутри той или иной серии Реферативного журнала.

Реферативная работа в ВИНИТИ организована таким образом, что ценная научная публикация, даже если она помещена в издании, профиль которого отличается от профиля данной серии Реферативного журнала, все же будет извлечена и прореферирована в этой серии. Это обеспечивается системой централизованной научной обработки информационных источников.

Принцип централизованной обработки мировой научно-технической литературы обусловлен тем, что между теоретическими и прикладными знаниями установилась органическая связь, в результате чего традиционное разделение науки и техники

на изолированные отрасли становится в значительной степени условным и искусственным.

В связи с этим с самого начала деятельности Института пришлось отказаться от журнального принципа распределения материалов, при котором за определенной отраслью науки или техники была бы закреплена определенная группа журналов. Было признано целесообразным обработку мировой научно-технической литературы производить на базе единого рубрикатора Реферативного журнала. Весь массив получаемой литературы просматривается в интересах каждой отрасли науки или техники, а выбор материала из этого массива определяется рубрикатором.

Другим основным информационным изданием ВИНИТИ является Экспресс-информация (ЭИ), представляющая собой периодическое издание, которое освещает наиболее актуальные вопросы науки и техники на основе зарубежных публикаций. В ЭИ используется форма сокращенных переводов статей, патентов и других научных документов; приводятся соответствующие иллюстрации, графики и чертежи. Как правило, ЭИ исключает необходимость обращения к оригиналу.

ЭИ предназначена для широких кругов инженерно-технических работников и рабочих высокой квалификации, а также для научных работников, преподавателей, аспирантов и студентов высших учебных заведений.

ЭИ издается в 50 сериях, посвященных важнейшим отраслям промышленности (например, имеются серии «Горнорудная промышленность», «Железнодорожный транспорт», «Нефтепромысловое дело», «Радиотехника и электроника», «Текстильная промышленность», «Химия и технология неорганических веществ», «Черная металлургия» и т. д.), а также по отдельным, наиболее актуальным научным проблемам (например, «Автоматическое управление производственными процессами», «Коррозия и защита металлов», «Синтетические высокополимерные материалы», «Вычислительная техника», «Астронавтика и ракетодинамика» и т. д.) [10]. Учитывая интерес широких кругов специалистов к теории, методике и практике научной информации, ВИНИТИ приступил к изданию ЭИ по теории и практике научной информации. Периодичность ЭИ — 48 выпусков в год.

В одном выпуске ЭИ обычно содержится пять—восемь статей и сообщений. Таким образом, в каждой серии за год дается около 300 вторичных публикаций. Каждая серия ЭИ имеет внутренние подразделы. Выпуск комплектуется в виде отдельных тетрадей по каждому подразделу. Это позволяет подписчику группировать материал по подразделам и хранить его в виде картотеки. Так,

[10] Полный перечень серий ЭИ см. в «Проспекте периодических изданий ВИНИТИ на 1968 г.» (М., 1967).

например, серия «Радиотехника и электроника» имеет подразделы: «Общий раздел», «Радиотехника», «Электроника», «Элементы и детали радиоаппаратуры», «Радиотехнические материалы», «Техника измерений и испытаний», «Телевидение» и т. д. Издание ЭИ осуществляется в три-четыре раза быстрее, чем издание Реферативного журнала.

В 1962 г. Институт приступил к подготовке нового вида информационного издания, получившего название «Итоги науки». Это издание предназначено для научного обобщения и систематизации достижений в отдельных отраслях науки и техники за определенный период — как правило, за календарный год — по материалам, освещенным в Реферативном журнале.

Такое обобщение облегчит специалисту ознакомление с результатами и основными направлениями развития данной отрасли науки или техники. Вместе с тем «Итоги науки» должны давать библиографию важнейших публикаций по соответствующей отрасли. Институт стремится систематически публиковать обзоры по важнейшим проблемам, причем каждый новый обзор как бы служит продолжением предыдущего. В качестве примера можно назвать следующие выпуски «Итогов науки»: «Химия и технология синтетических высокомолекулярных соединений», «Химия, металловедение и обработка титана», «Успехи в изучении океанических глубин» и т. д.[11]

Все принципиальные научные и методологические вопросы, возникающие при подготовке информационных изданий Института, решает Главная редколлегия, которая является совещательным органом при директоре. В состав Главной редколлегии входят директор, заместители директора, ученый секретарь, главные редакторы отдельных серий Реферативного журнала и персонально приглашенные специалисты. Главная редколлегия обеспечивает правильную идеологическую и научную направленность информационных изданий Института; утверждает рубрикаторы этих изданий, ежегодные планы издания научно-информационной литературы; заслушивает отчеты о научно-редакционной деятельности главных редакторов; обсуждает вопросы, связанные с научно-методической подготовкой и обменом информационных материалов, а также их распространением и пропагандой.

Для непосредственного руководства подготовкой к изданию сводных томов и отдельных выпусков Реферативного журнала, а также других видов информационных изданий создаются редакционные коллегии в отраслевых отделах. Организационная структура ВИНИТИ показана на рис. 119.

---

[11] Перечень выпускаемых ВИНИТИ обзоров ежегодно публикуется в издании: «Проспект „Итоги науки“ и „Итоги науки и техники“» (по материалам за. . . год).

наибольшие успехи достигнуты в разработке информационно-поисковых языков для химии и электротехники. Большое значение имеют работы по сравнению эффективности научно-информационной деятельности вообще и информационно-поисковых систем в частности: установление критериев сравнения, проблемы релевантности документов, сравнительные экспериментальные исследования и т. д.

4. Разработка средств технической реализации информационных систем. Основой автоматической информационной системы является универсальная ЭВЦМ. Однако эффективное использование ЭВЦМ для решения научно-информационных задач возможно лишь при наличии специальных устройств ввода и вывода буквенно-цифровой информации, а также специальных запоминающих устройств. В ВИНИТИ ведутся работы по всем этим направлениям. Разрабатываются читающие автоматы, приспособленные для непосредственного оптического ввода текста в машину. Ведется также работа по созданию быстродействующих устройств, предназначенных для вывода из машины буквенно-цифровой информации в любом нужном алфавите, а также различных графов и кривых. В целях существенного уменьшения времени поиска, особенно при больших массивах хранящейся информации, ведутся работы по созданию запоминающих устройств ассоциативного типа. В этих устройствах поиск необходимых сведений ведется не по адресу ячейки, в которой эти сведения записаны, а непосредственно по признакам, характерным для записанных данных.

Основная задача ВИНИТИ — обеспечение ученых и специалистов исчерпывающей информацией по всем публикуемым научным документам. При решении этой задачи ВИНИТИ исходит из следующих основных положений:

1. Аналитико-синтетическая переработка каждого первичного документа, поступающего в Институт, должна производиться только один раз, но выполняться таким образом, чтобы данные, полученные в результате такой переработки, позволяли обеспечивать самые различные виды информационного обслуживания — выпускать сигнальные издания и реферативные журналы, подготавливать различные указатели к реферативным журналам, вводить документы в автоматическую информационно-поисковую систему и т. д.

2. Централизация аналитико-синтетической переработки всей научно-технической литературы в ВИНИТИ создает хорошие условия для предельной механизации и автоматизации такой переработки.

3. Создание в ВИНИТИ автоматической информационно-поисковой системы (например, типа «Медларс») позволяет обеспечить быстрый выпуск сигнальных изданий, а также производить информационный поиск по любым запросам, поступающим извне.

Целесообразно создание сети поисковых систем такого же типа в различных пунктах Советского Союза. Эти поисковые системы должны снабжаться магнитными лентами с соответствующей информацией, которые подготавливаются в центральной поисковой системе Института.

4. Интегральным элементом системы научно-информационного обслуживания, создаваемой на базе ВИНИТИ, является служба изготовления копий научных документов по запросам потребителей. Поток таких запросов на фотокопии представляет собой важный элемент обратной связи Института с потребителями, позволяющей постоянно оптимизировать работу всей информационной системы.

5. Создание на базе ВИНИТИ такой системы информационного обслуживания требует осуществления множества целенаправленных научно-исследовательских и опытно-конструкторских работ, а отладка и эксплуатация всей системы позволяет накапливать ценнейшие фактические сведения, без которых невозможна плодотворная теоретическая работа в области информатики.

6. Для оказания помощи потребителям в отыскании нужной научной информации в ВИНИТИ создается всесоюзный справочный центр, в котором будут сосредоточены сведения о всех информационных центрах страны, располагающих специализированными фондами научных документов (как опубликованных, так и неопубликованных). В этом центре должны накапливаться сведения не только о крупных отраслевых фондах (их тематика, размещение, порядок запроса и т. п.), но и о фондах отдельных предприятий, институтов и лабораторий по узким вопросам.

## Всесоюзный научно-технический информационный центр (ВНТИЦ)

Всесоюзный научно-технический информационный центр (ВНТИЦ) с общегосударственным фондом микрокопий отчетов об этих работах создан постановлением Совета Министров СССР № 916 от 29.XI 1966 г. «Об общегосударственной системе научно-технической информации».

Необходимость создания такого всесоюзного центра диктовалась следующими соображениями. В настоящее время в СССР насчитывается несколько тысяч научно-исследовательских учреждений различных направлений, конструкторских бюро и проектных организаций, которые ежегодно выполняют десятки тысяч научно-исследовательских и опытно-конструкторских работ. Однако в стране до сих пор не организована систематическая информация о результатах этих работ. В периодической литературе освещается не более 25—30% общего числа исследований. Между тем достоверная и полная информация о проводимых

что центрами патентных фондов будут головные отраслевые научно-исследовательские институты, где патентные материалы найдут отражение в соответствующих справочно-информационных фондах.

Издаваемые ЦНИИПИ материалы об отечественных и зарубежных изобретениях являются важнейшим источником информации для ученых, конструкторов, технологов, изобретателей, рационализаторов и новаторов производства, а также для работников патентных служб, связанных с определением патентоспособности и патентной «частоты» промышленных изделий. ЦНИИПИ выпускает следующие основные виды информационных изданий по патентам.

1. «Табуляграммы описаний изобретений». Табуляграммы издаются по разным отраслям техники и включают библиографические сведения об отечественных и зарубежных изобретениях: номер и дату выдачи авторского свидетельства или патента, страну, в которой выдан патент, название фирмы, являющейся его владельцем, порядковый номер кадра на микрофильме с описанием изобретений к патентам.

2. Бюллетень «Изобретения, промышленные образцы и товарные знаки».

3. Бюллетень «Информация по изобретательству».

4. Описания изобретений к авторским свидетельствам и патентам, выдаваемым в СССР.

5. Перфокарты с переводом публикаций об изобретениях из патентных бюллетеней США, Великобритании, ФРГ.

6. Технико-экономические обзоры изобретений по отраслям народного хозяйства.

7. Библиографические сборники патентных материалов стран с небольшим патентным фондом.

8. Сведения об описаниях отечественных изобретений (на библиотечных карточках).

9. Методические материалы по патентной классификации.

10. Сведения о новых поступлениях описаний зарубежных изобретений во Всесоюзную патентно-техническую библиотеку (на библиотечных карточках).

### Всесоюзный научно-исследовательский институт научно-технической информации, классификации и кодирования (ВНИИКИ)

В системе Комитета стандартов, мер и измерительных приборов при Совете Министров СССР созданы Всесоюзный научно-исследовательский институт научно-технической информации, классификации и кодирования (ВНИИКИ) и Всесоюзный информационный фонд стандартов и технических условий (ВИФС), которые выпускают информационные материалы по вопросам

стандартизации, повышения качества, надежности и долговечности продукции, теоретической и практической метрологии, измерительной техники, классификации, кодирования и терминологии в нашей стране и за рубежом.

Эти материалы должны способствовать решению основных задач по обеспечению выпуска продукции высокого качества; повышению роли государственных стандартов в управлении качеством промышленной продукции; повышению надежности и долговечности промышленной продукции путем широкого внедрения методов унификации; созданию единой системы нормативной и технической документации; повышению уровня техники измерений во всех областях науки и производства.

Органы информации Комитета стандартов, мер и измерительных приборов при Совете Министров СССР издают библиографическую, реферативную и обзорную информацию.

### *Отраслевые системы научно-технической информации*

Как мы видели, усилиями всесоюзных органов научно-технической информации обеспечивается централизованная обработка опубликованных научно-технических документов. Всего 15—20 лет назад этого, вероятно, было бы достаточно для того, чтобы полностью обеспечить потребителя библиографической и реферативной информацией по интересующей его тематике. В то время сведения о каждом сколько-нибудь важном открытии или достижении публиковались, а темпы развития науки и техники были такими, что можно было не принимать во внимание время, проходившее от момента этого открытия до публикации сведений о нем в каком-нибудь издании.

В настоящее время на заводах, в научно-исследовательских и проектно-конструкторских организациях создается очень большое количество информационных материалов — отчетной, проектной, технологической документации. Далеко не вся эта информация попадает в печать. И что самое главное — неопубликованной остается отнюдь не второстепенная информация. Сейчас не существует порядка, который гарантировал бы отбор и обязательную публикацию наиболее ценных материалов. Такой порядок нелегко создать и, по-видимому, очень трудно поддерживать. Поэтому очень много ценной информации остается там, где она была создана — на заводах, в НИИ, проектных институтах и конструкторских бюро. В массе неопубликованной информации остается большинство так называемых частных решений, полученных в ходе научно-исследовательских и опытно-конструкторских работ. Под частными решениями понимают самостоятельные, иногда не имеющие прямого отношения к теме исследования, результаты. Среди них могут быть очень важные для других работ методики

Ниже в качестве примера даются отдельные сведения о деятельности некоторых центральных отраслевых органов научно-технической информации. При этом следует иметь в виду, что первые три из них (ВНИИМИ, ВИНТИСХ и ЦИНИС) частично выполняют функции всесоюзных органов научно-технической информации, выпуская реферативные издания в дополнение к РЖ ВИНИТИ.

*Всесоюзный научно-исследовательский институт медицинской и медико-технической информации* (*ВНИИМИ*) является центральным отраслевым информационным органом в области медицинской и медико-технической информации. Институт организует и координирует работу отделов научно-технической информации в республиках и информационных отделов и групп в научных институтах; осуществляет методическое руководство работой служб научно-технической информации в системе здравоохранения; организует научно-информационное обслуживание на основе неопубликованных источников и обеспечивает централизованную обработку опубликованных работ в области медицины и здравоохранения. Для подготовки своих публикаций ВНИИМИ использует обширный отечественный и зарубежный материал — более 2000 наименований периодических и продолжающихся изданий на 40 языках мира.

С целью ознакомления научно-исследовательских и лечебно-профилактических учреждений, а также широких кругов медицинской общественности с достижениями отечественной и зарубежной медицинской науки и техники ВНИИМИ издает информационные издания по вопросам медицины и здравоохранения. Одним из основных изданий Института является ежемесячный «Медицинский реферативный журнал», который дает информацию по 38 медицинским дисциплинам. На основе материалов РЖ по отдельным проблемам медицины издается реферативная картотека, которая может быть использована для создания справочного информационного фонда в органах научной информации. Текущая реферативная информация содержится в информационных письмах (24 выпуска в год объемом по 0,5 авт. л. каждый). Они знакомят читателей с достижениями медицины за рубежом, с новыми методами лечения, новыми препаратами, статистическими и демографическими данными и т. д. Институт выпускает аннотированный каталог переводов статей из иностранных медицинских журналов, сделанных в научных учреждениях системы здравоохранения, а также публикует информацию о всесоюзных и республиканских медицинских съездах и конференциях, международных, региональных, национальных конгрессах, съездах, конференциях, симпозиумах и коллоквиумах.

ВНИИМИ издает реферативные сборники («Медицина за рубежом»). Уже изданы: «Вопросы онкологии», «Заболевания сердечно-сосудистой системы», «Гигиена воды, воздуха и почвы»,

«Вопросы организации здравоохранения и истории медицины», «Медицинская техника», «Вопросы фармации и аптечного дела», «Радиоэлектроника в медицине» и др. В них печатаются рефераты монографий и статей, посвященных актуальным проблемам медицинской науки, медицинской техники, организации здравоохранения и т. п. Институт подготавливает и издает научные обзоры. Так, например, были изданы обзоры: «Физиология и патология сердечно-сосудистой системы»; «Вирусы и вирусные заболевания»; «Злокачественные новообразования»; «Гипертоническая болезнь, атеросклероз и коронарная недостаточность»; «Вопросы организации здравоохранения и история медицины»; «Медицинская техника» и др. ВНИИМИ, кроме того, осуществляет библиографическую информацию о достижениях медицины за рубежом. Библиографические картотеки издаются по разделам медицинской науки и техники, а также по законченным диссертационным работам.

*Всесоюзный институт научно-технической информации по сельскому хозяйству* (*ВИНТИСХ*) выполняет следующие функции:

1) создание системы органов научно-технической информации по отрасли и определение форм и содержания их работы;

2) организацию научно-технической информации по опубликованным и неопубликованным источникам;

3) методическую работу и ее координацию с органами информации других отраслевых научно-исследовательских учреждений;

4) справочно-информационное обслуживание по запросам руководящих партийных, советских и сельскохозяйственных органов, отдельных научных и опытных учреждений, а также и отдельных специалистов.

В соответствии с поставленными задачами ВИНТИСХ выпускает следующие типы изданий:

1. Указатель законченных в СССР научно-исследовательских работ, рекомендуемых для внедрения в сельскохозяйственное производство. (По материалам отчетов научно-исследовательских институтов и опытных станций).

2. Новости сельскохозяйственной науки и практики.

3. Сельскохозяйственная наука и практика за рубежом.

4. Реферативная картотека отечественной и иностранной литературы.

5. Обзоры отечественной и зарубежной литературы, сборники переводов по важнейшим проблемам сельскохозяйственной науки и передовому опыту.

6. Справочники.

*Центральный институт научной информации по строительству и архитектуре* (*ЦИНИС*) по положению является центральным отраслевым, однако он, как и два предыдущих института,

*Научно-исследовательский институт экономики и информации по радиоэлектронике* (*НИИЭИР*) является центральным в своей отрасли информационным органом. Его основные функции: подготовка реферативной информации по ведомственным источникам информации и непубликуемым источникам в целях централизованного пополнения местных фондов и их справочных аппаратов, создание и пополнение отраслевого фонда, оперативное информирование о новых направлениях развития техники, составление обзоров по общетехническим вопросам, ответы на запросы предприятий, методическое руководство отраслевой сетью информационных органов, научно-исследовательские работы по совершенствованию информационной системы и ее технических средств [13].

Отраслевой информационный центр по радиоэлектронике функционирует с 1954 г. К настоящему времени единая система научно-технической информации по радиоэлектронике складывается из следующих взаимосвязанных систем: потоков информации, обработки информации, документальной информации, справочно-информационных фондов, информационного обслуживания, устной и наглядной информации, научно-исследовательских работ по совершенствованию информационной системы, управления информационными органами.

НИИЭИР разрабатывает проекты создания единой системы построения информационных органов в отрасли и потоков информации между ними, системы отбора информации из отчетной научно-технической документации. В Институте разработан ряд информационно-поисковых систем: ИПС на перфокартах ручного обращения, ИПС на суперпозиционных перфокартах большой емкости (до 22 тыс. документов в каждом подмассиве), ИПС, реализованная на ЭЦВМ «Минск 22», осуществляющая поиск из ретроспективного фонда и выдачу с алфавитно-шифрового печатающего устройства инвентарных номеров искомых рефератов и библиографических списков по заданной теме.

До настоящего времени научно-информационная деятельность в нашей стране развивалась главным образом в области естественных и прикладных наук. Теперь назрела необходимость повысить эффективность научно-информационной деятельности *в общественных науках*. В августе 1967 г. ЦК КПСС принял постановление «О мерах по дальнейшему развитию общественных наук и повышению их роли в коммунистическом строительстве», в котором особо отмечена важность проблемы научной информации.

В постановлении, в частности, сказано: «В целях коренного улучшения системы научной информации и дальнейшего расшире-

[13] А. Г а у х м а н. Организация информационной работы в области радиоэлектроники. «Научно-техн. информация», 1967, серия 1, № 11, 13—20.

ния документальной базы общественных наук считать необходимым: а) Государственному комитету Совета Министров СССР по науке и технике, президиуму Академии наук СССР и Министерству высшего и среднего специального образования СССР принять меры по увеличению объема и обеспечению систематического поступления информационных материалов, сокращению сроков обработки и повышению качественного уровня информации; сделать ее достоянием широкого круга научных работников и преподавателей; подготовить предложения об организации Института научной информации по общественным наукам»[14].

### *Территориальные органы научно-технической информации*

В общегосударственной системе научно-технической информации особое место занимают институты научно-технической информации в союзных республиках и территориальные центры научно-технической информации и пропаганды.

В соответствии с Постановлением Совета Министров СССР № 916 от 29.XI 1966 г. на Советы Министров союзных республик, а по РСФСР — на Государственный Комитет Совета Министров СССР по науке и технике возложена ответственность за развитие межотраслевых территориальных органов информации.

На институты научно-технической информации в союзных республиках возложено:

1) оперативное и дифференцированное обеспечение научно-технической информацией правительственных органов республик;

2) организация справочно-информационного обслуживания предприятий и организаций, расположенных на территории республики, в том числе обеспечение по их запросам научно-технической литературой, патентными материалами, стандартами и др.;

3) распространение в республике научно-технической информации и передового производственного опыта по межотраслевым проблемам;

4) подготовка и издание информационных материалов на национальных языках в соответствии с планом, утверждаемым Государственным комитетом по науке и технике;

5) контроль за представлением в установленном порядке предприятиями и организациями республиканского подчинения информационных сообщений и научно-технической документации в соответствующие отраслевые информационные центры;

6) разработка и внедрение механизированных и автоматизированных систем подготовки и поиска научно-технической

[14] О мерах по дальнейшему развитию общественных наук и повышению их роли в коммунистическом строительстве. Постановление ЦК КПСС. «Коммунист», 1967, № 13, стр. 111

2) справочно-информационное обслуживание предприятий и организаций, расположенных на данной территории, в особенности предприятий и организаций местного подчинения, и не имеющих собственных информационных подразделений;

3) распространение научно-технической информации и передового производственного опыта по межотраслевым вопросам;

4) проведение научно-технической пропаганды;

5) оказание методической помощи информационным подразделениям предприятий и организаций.

На рис. 122 приведена схема построения сети органов научно-технической информации в РСФСР, утвержденная Государственным комитетом Совета Министров СССР по науке и технике в 1967 г.

### *Роль информационно-поисковых систем в общегосударственной системе научно-технической информации*

Накопление большого числа документов в отдельных отраслях промышленности, на некоторых промышленных предприятиях, в научно-исследовательских и проектных институтах привело к необходимости создания справочно-информационных фондов с системой машинного или ручного поиска документов.

Разработка информационно-поисковых систем в настоящее время приняла довольно широкий размах в связи с резким увеличением сведений по всем отраслям науки и техники и возросшими требованиями потребителей информации к срокам информационного обслуживания, а также ускорением темпов технического прогресса. Еще несколько лет назад потребители информации могли быть удовлетворены поиском материалов при помощи только простейших ИПС, вплоть до немеханизированных библиотечно-библиографических ИПС.

Теперь такие системы уже не могут удовлетворить всех запросов научных и инженерно-технических работников. Объемы справочно-информационных фондов значительно возросли, все больше и большее значение приобретают межотраслевые проблемы, представленные многоаспектными материалами. В связи с этим усложнились требования к системам поиска, которые используются теперь для подбора материалов из огромных массивов информации с гораздо более точным, чем ранее, учетом различных смысловых аспектов этих материалов. Однако это весьма важное для разработки всей проблематики ИПС принципиальное положение получило широкое признание лишь за последние несколько лет.

Большое число опубликованных работ по этой проблеме, а также непосредственный опыт ряда исследовательских организаций у нас в стране дают достаточное представление о значении

проблемы и достигнутых результатах. Принципиальная и техническая стороны вопроса построения ИПС достаточно полно изложены нами в главах 4—7. Здесь мы касаемся лишь научно-организационного аспекта этой проблемы. Идея государственного справочно-информационного фонда разрешается путем создания ИПС в масштабе государства на основе взаимосвязанных справочно-информационных фондов информационных органов разных уровней. Эта работа должна вестись по единому плану под непосредственным государственным контролем. Справочно-информационные фонды всесоюзных и отраслевых информационных центров должны функционировать как части единого целого, образуя государственный справочно-информационный фонд по естественным и техническим наукам.

В связи с этим большое значение приобретает четкое распределение обязанностей между отдельными частями этой единой системы как по линии оперативного обслуживания потребителей, так и по линии научно-исследовательских работ. Рассмотрим вначале организационные вопросы обслуживания, так как именно практика формирует тематику теоретических и экспериментальных работ. Основным звеном системы общегосударственного справочно-информационного фонда является отраслевой фонд, представляющий собой совокупность собственного фонда отраслевого института информации и фондов научно-исследовательских институтов, проектно-конструкторских организаций и предприятий соответствующей отрасли. Этот отраслевой фонд имеет единый справочный аппарат, сосредоточенный, как правило, в отраслевом институте информации и обеспечивающий возможность ответа на запрос по материалам всех частей фонда, независимо от их места хранения.

Отраслевой фонд пополняется опубликованными материалами из всесоюзных институтов информации (по периодическим и продолжающимся изданиям, трудам съездов, конференций, симпозиумов, монографиям и т. п. — из ВИНИТИ, по патентам — из ЦНИИПИ, по стандартам — из ВНИИКИ) и неопубликуемыми материалами — из организаций и предприятий отрасли (отчеты о законченных научно-исследовательских и опытно-конструкторских работах или их крупных этапах, обзоры научно-технических достижений в отрасли, материалы по обмену опытом и т. п.). В этих же фондах депонируются рукописи по тематике данной отрасли, размножение которых массовыми тиражами признано нецелесообразным.

За последние годы в большинстве отраслевых информационных центров страны созданы крупные справочно-информационные фонды. Например, в институтах информации радиопромышленности собрано по 700 тыс. документов, в станкостроительной и инструментальной промышленности — более 100 тыс. документов. На основе этих фондов удовлетворяются запросы предприятий и

организаций, отдельных ученых и специалистов. Так, Институт информации по электротехнике в 1965 г. ответил более чем на 300 тыс. запросов, Институт информации по радиоэлектронике — на 130 тыс. запросов и т. д.

Важное значение имели решения о прекращении выпуска изданий журнального типа отраслевыми информационными центрами и об организации издания информационных материалов на картах, предназначенных для создания поискового аппарата, справочно-информационных фондов предприятий и организаций. Это существенно облегчит комплектование справочно-информационных фондов на предприятиях и в организациях, позволит ускорить доведение до специалистов необходимой им информации. На основе таких карточек в институтах, на крупных и средних предприятиях создаются местные справочно-информационные фонды, учитывающие профиль деятельности предприятия (института).

Наконец, следует отметить вид деятельности, который проводится всеми институтами информации, независимо от их положения. Это изготовление копий научно-технических документов по запросам потребителей. Необходимо подчеркнуть, что для ученого и инженера основным источником научно-технической информации является первичный документ — статья, отчет, патентное описание, монография, схема, чертеж и т. п. Поэтому должна быть создана такая система, при которой любой ученый, инженер и новатор производства сможет быстро получить любой нужный документ или его копию. О размахе этой деятельности можно судить по следующим данным: ВИНИТИ в 1966 г. выслал по запросам институтов, предприятий и отдельных специалистов более 1 млн. страниц фото- и электрографических копий публикаций и 0,5 млн. кадров микрофильмов; большое число копий оригинальных материалов разослали также отраслевые институты информации.

Картина справочно-информационного обслуживания будет неполной, если мы не упомянем о роли территориальных органов информации: республиканских институтов информации и ЦБТИ экономических районов РСФСР. Территориальные органы информации могут снять значительную часть нагрузки в информационно-справочной деятельности всесоюзных и отраслевых органов информации путем дублирования той части всесоюзного информационно-справочного фонда, которая охватывает характерные для данной республики (экономического района РСФСР) отрасли науки и техники. Кроме того, республиканские институты информации должны обеспечивать возможность получения запроса и ответа на него на национальном языке республики.

Задачу создания в стране высокоэффективной справочно-информационной службы можно решить только на основе широ-

ких теоретических и экспериментальных исследований. Задачи резкого повышения оперативности и качества справочно-информационного обслуживания требуют оснащения информационных органов современными техническими средствами для механизации и автоматизации процессов поиска информации, а также копирования, размножения документов и передачи их на расстояние.

Всесоюзные и отраслевые информационные центры должны обеспечить выпуск информационных материалов на унифицированных носителях — перфокартах, микрокартах, перфолентах, магнитных лентах, что позволит организовать быстрый поиск информации на местах. Стандартные носители информации позволят применить для технической реализации ИПС в самых широких размерах ЭЦВМ, так как стандартные носители создадут возможность автоматического ввода сведений в ЭЦВМ без дополнительной подготовки и переноса информации с одного носителя на другой, что значительно ускорит организацию поиска и улучшит экономические показатели системы.

## *Роль научно-технических библиотек в общегосударственной системе научно-технической информации*

По общности предмета исследования и некоторых этапов практической деятельности информатика и научно-информационная деятельность связаны с библиотековедением, библиотечным делом и библиографией. Более того, до возникновения информатики как самостоятельной дисциплины некоторые ее задачи решались именно в этих областях. Информатика складывалась в значительной степени вне названных традиционных отраслей человеческой деятельности. Вместе с тем библиотеки и библиографические учреждения продолжают вести информационную работу. Это породило ряд неверных представлений о взаимосвязях библиотечно-библиографической и научно-информационной деятельности.

Мы исходим из того, что советская библиотека — это культурно-просветительное и научное учреждение, организующее общественное пользование произведениями печати и письменности и призванное помогать читателю в выборе книг, руководить чтением для того, чтобы широко пропагандировать идеи коммунизма, поднимать идейно-политический и культурный уровень советского народа, мобилизовать его на решение политических, научных и хозяйственных задач. Говоря о библиотеках, необходимо подчеркнуть специфические задачи научных и специальных библиотек. Основной задачей этих библиотек является не общепросветительная работа, а оказание помощи развитию науки и техники. В последние годы осуществлен ряд мероприятий, направ-



1/4 45 Зак. № 864

ленных на обеспечение единства деятельности научно-технических библиотек и информационных органов. В результате содержание работы научно-технических библиотек обогатилось за счет включения в их фонды различных непубликуемых материалов (отчеты по научно-исследовательским работам, непубликуемые обзоры, депонированные рукописи и др.). Организация совместной работы научно-технических библиотек и информационных органов позволила использовать в общих интересах более мощную техническую базу, имеющуюся в информационных органах, а также привлечь к комплектованию и анализу научно-технической литературы квалифицированных специалистов, работающих в информационных органах.

Постановлением Совета Министров СССР № 755 от 10.IX 1964 г. «Об улучшении научно-технической информации в стране» установлено, что научно-технические библиотеки являются важной составной частью системы научно-технической информации. На научно-технические библиотеки в организациях и на предприятиях, где нет информационных подразделений, возложена ответственность за организацию справочно-информационной работы. В связи с этим необходимо последовательно претворять в жизнь этот принцип и до конца преодолеть мешающий дальнейшему развитию параллелизм в деятельности научно-технических библиотек и информационных органов. Научно-технические библиотеки (республиканские, отраслевые, на предприятиях и в организациях) должны быть включены в качестве самостоятельных структурных подразделений в состав соответствующих информационных органов. Деятельность указанных библиотек в этом случае еще в большей степени будет направлена на обеспечение научных специалистов и рабочих-новаторов необходимой им литературой и различными информационными материалами о достижениях науки и техники, а также о передовом производственном опыте.

Постановление Совета Министров СССР № 916 от 29.XI 1966 г. «О единой государственной системе научно-технической информации» внесло существенное размежевание в задачи и функции научно-технических библиотек. Этим постановлением Министерство культуры СССР возложено совершенствование научно-информационной работы универсальных библиотек подведомственных Министерству культуры, в первую очередь Государственной библиотеки им. В. И. Ленина, Государственной библиотеки им. М. Е. Салтыкова-Щедрина, Всесоюзной государственной библиотеки иностранной литературы, а также республиканских, краевых и областных библиотек, имея в виду использование фондов этих библиотек в общегосударственной системе справочно-информационной службы, координацию работы с деятельностью органов научно-технической информации и специальных библиотек.

В постановлении определяется самостоятельное место научно-технических библиотек как одного из основных звеньев в общегосударственной системе научно-технической информации. Общее научное и методическое руководство и координация деятельности специальных, научных и научно-технических библиотек возлагается на Межведомственную комиссию при Государственном Комитете СССР по науке и технике.

Другой аспект рассматриваемого вопроса — место библиотек и библиографических учреждений в общей системе научной информации. Он затрагивает ряд организационных моментов. Используя весь арсенал средств работы с читателем, библиотека может оказывать действенную помощь в осуществлении обратной связи с потребителями информации. Необходимость в этом очень велика, так как обратная связь позволяет установить контроль за действенностью информационных изданий. Поскольку научно-информационная деятельность в настоящее время базируется на литературных источниках, очень важно выделить наиболее ценные в научном отношении книжные и периодические издания. Организация тщательного изучения мирового книжного и журнального фондов может послужить важным научным вкладом библиотековедения.

Правильное сочетание задач и функций научно-информационной деятельности и библиотечного обслуживания имеет и ряд других аспектов, например, такой, как информационная деятельность библиотек и библиографических учреждений. Справочно-библиографическое и информационное обслуживание продолжает оставаться важной функцией специальных библиотек. С развитием информатики эта деятельность может быть значительно улучшена. Справочно-библиографическая служба библиотек становится более эффективной благодаря использованию реферативных журналов. С расширением издания картотек и бюллетеней экспресс-информации отпадает необходимость ведения этой работы в библиотеках. Это высвобождает время и силы библиотечного персонала для более углубленной работы по подготовке тематических ретроспективных справок, созданию отраслевых библиографических указателей. Применение некоторых методов информатики, например, координатного индексирования документов, поможет сделать эту работу более эффективной и использовать для ее выполнения новейшие технические средства. Эти соображения справедливы и для таких традиционных библиотечных процессов, как каталогизация и информация о новых поступлениях в библиотеку, обслуживание читателей и др.

## НАУЧНО-ИНФОРМАЦИОННАЯ ДЕЯТЕЛЬНОСТЬ В СТРАНАХ — ЧЛЕНАХ СЭВ

В организации научной и технической информации в социалистических странах отражены черты их государственного строя и принципа государственного планирования и управления народным хозяйством.

Несмотря на то, что в каждой социалистической стране имеются определенные исторически сложившиеся экономические условия и в связи с этим к научно-технической информации предъявляются различные требования, общим для всех этих стран является государственная координация информационной деятельности. Руководство научно-информационной деятельностью осуществляется соответствующими центральными государственными органами через подчиненные им национальные информационные центры.

Такими центральными информационными учреждениями являются:

в Народной Республике Болгарии — Центральный институт научно-технической информации;

в Венгерской Народной Республике — Центр технической документации при Венгерской центральной технической библиотеке;

в Германской Демократической Республике — Центральный институт информации и документации;

в Польской Народной Республике — Центральный институт научно-технической и экономической информации;

в Социалистической Республике Румынии — Центральный институт технической документации;

в Чехословацкой Социалистической Республике — Центр научной, технической и экономической информации.

На центральные органы научной и технической информации в социалистических странах возложены следующие задачи:

1) осуществление централизованной информационной деятельности в области научных исследований и техники, экономических исследований, а также по вопросам организации труда, организации производства и управления;

2) рассмотрение планов проведения научно-исследовательских работ в области информационной деятельности;

3) руководство техническими библиотеками;

4) изучение вопросов эффективности использования научной информации;

5) разработка рекомендаций по организации сети информационных органов в стране, а также методическое руководство деятельностью отраслевых центров информации;

6) организация подготовки и усовершенствования кадров органов научно-технической информации;

7) сотрудничество с научными, техническими, производственными и общественными национальными и международными организациями.

Следующей ступенью информационной деятельности в социалистических странах являются отраслевые центры. Задачи, возлагаемые на отраслевые информационные центры, сводятся к следующему:

1) обеспечение исчерпывающей информацией по тематике данной отрасли всех заинтересованных организаций и отдельных лиц;

2) оказание помощи в организации отделов информации на предприятиях данной отрасли.

В отраслевых центрах социалистических стран широко практикуется составление и издание информационных карточек. В ряде социалистических стран информационные карточки являются основной формой информации. Карточки рассылаются по подписке предприятиям, научным учреждениям и индивидуальным подписчикам.

Низовую сеть информационной деятельности в социалистических странах составляют отделы информации на промышленных предприятиях и в научных учреждениях. В задачи этих низовых звеньев входит:

1) обеспечение научно-технической информацией данного предприятия;

2) развитие всех форм технико-экономической пропаганды;

3) наблюдение за деятельностью технических библиотек.

Отделы информации на предприятиях обычно сами не издают информационных материалов, а используют информационные материалы, издаваемые отраслевыми центрами научно-технической информации.

Таковы в общих чертах основные организационные принципы, на которых строится научно-информационная деятельность в социалистических странах. Далее в краткой форме описаны системы и организация информационной деятельности в отдельных странах.

### *Народная Республика Болгария*

Научно-информационная деятельность в Народной Республике Болгарии (НРБ) организована и получила широкое развитие за годы народной власти. Постановлением Правительства НРБ от 1962 г. научно-техническая информация в НРБ получила дальнейшее развитие и организована в единую государственную систему. Общее государственное руководство в стране научно-информационной деятельностью возложено на Государственный комитет по науке и техническому прогрессу. Структура и организация научно-информационной деятельности в НРБ создана в соответствии с государственным руководством в стране, научно-исследовательской, проектно-конструкторской и производственной

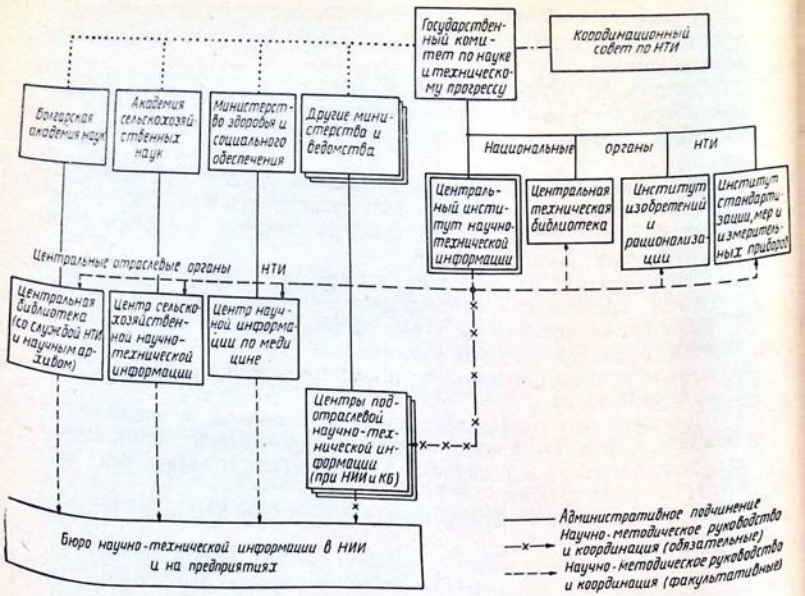



**Рис. 123. Организация научно-информационной деятельности в Народной Республике Болгарии**

деятельностью. На рис. 123 приведена схема единой государственной системы научно-технической информационной деятельности НРБ. Эта система состоит из следующих основных звеньев.

I. Национальные информационные органы.
II. Отраслевые информационные органы.
III. Подотраслевые информационные органы.
IV. Низовые информационные органы.

Государственному Комитету по науке и техническому прогрессу непосредственно подчинены следующие *национальные информационные центры.*

1) Центральный институт научно-технической информации;
2) Центральная техническая библиотека;
3) Институт изобретений и рационализации;
4) Институт стандартизации, мер и измерительных приборов.

Центральный институт научно-технической информации является главным информационным органом в единой государственной системе научно-технической информации. Он осуществляет научно-методическое руководство и координацию информационной

деятельности в стране. Институт издает обзорную информацию по межотраслевым и некоторым отраслевым вопросам в области промышленности, строительства и транспорта, выпускает на иностранных языках реферативную информацию болгарских научных и технических публикаций; подготавливает текущую информацию о законченных научно-исследовательских темах, об отчетах специалистов о заграничных командировках и материалах, получаемых в результате экономического и научно-технического сотрудничества; координирует переводы научно-технических публикаций и издает периодический бюллетень выполненных переводов; организует курсы по подготовке и переподготовке информационных кадров.

Центральная техническая библиотека осуществляет сбор и комплектование справочно-информационного фонда общегосударственного значения в области промышленности, строительства и транспорта.

Институт изобретений и рационализации проводит сбор и комплектование национального фонда патентной литературы и осуществляет справочно-библиографическую деятельность.

Институт стандартизации, мер и измерительных приборов организует сбор болгарских государственных стандартов, а также мировых стандартов и подготавливает и издает библиографическую информацию о новых и измененных болгарских государственных стандартах.

Центральная библиотека Болгарской академии наук подготавливает реферативную информацию по естественным и общественным наукам и издает бюллетень на иностранных языках о болгарских публикациях.

Центр сельскохозяйственной научно-технической информации проводит информационную деятельность в области сельского и лесного хозяйства и издает обзорную и реферативную информацию.

Центр научной медицинской информации осуществляет реферативную и обзорную информацию в области медицины и фармакологии.

*Подотраслевыми и информационными органами* являются центры подотраслевой научно-технической информации. Эти центры созданы при ведомственных научно-исследовательских и проектно-конструкторских институтах. Они комплектуют справочно-информационный фонд в зависимости от профиля соответствующего института; систематически обрабатывают комплектующийся справочно-информационный фонд и обеспечивают необходимой информацией научно-исследовательскую и проектно-конструкторскую деятельность; изготовляют первичную информацию о законченных темах и заданиях; снабжают информацией соответствующие низовые информационные органы, научно-методически руководят ими и координируют их деятельность.

*Низовыми информационными органами* являются бюро научно-технической информации, которые создаются в проектных организациях, на крупных предприятиях, в институтах академии и т. д. Деятельность этих бюро направлена на практическое использование информации. Они комплектуют также справочно-информационный фонд, который имеет узкую тематику в пределах интересов данного учреждения. Бюро научно-технической информации изготовляют первичную информацию о новых изделиях и передовом производственном опыте и законченных исследовательских темах и заданиях.

Различные органы научно-технической информации оперативно связаны между собой и ведут взаимный обмен информационными материалами. Единая государственная система научно-технической информации НРБ функционирует уже несколько лет. Практика показывает, что эта система очень жизненна и способствует научному и техническому прогрессу в стране. Осуществление единой государственной научно-технической информационной деятельности в НРБ подтвердило правильность принципов организации ее системы.

### *Венгерская Народная Республика*

Организация государственной службы научно-технической информации в Венгерской Народной Республике (ВНР) началась в 1948 г. На первом этапе информационное обслуживание велось главным образом техническими библиотеками предприятий и научных учреждений. В дальнейшем быстрый темп развития производства и научно-исследовательской деятельности выдвинул новые задачи перед информационным обслуживанием и потребовал расширения сети информационных органов. Поэтому в 1962 г. постановлением правительства был создан Государственный комитет по развитию техники, и все функции по руководству научно-информационной деятельностью и по технической пропаганде были переданы этому комитету. Круг задач, который был поставлен перед вновь созданным комитетом, сводится к следующему:

1) всемерно содействовать развитию сети научно-технической информации, устанавливая основные принципы этого развития;

2) координировать деятельность научно-информационных учреждений;

3) принимать участие в создании новых информационных органов, разрабатывая для этого положения и нормативы;

4) осуществлять контроль за подготовкой кадров в области научно-технической информации;

5) обеспечивать международное сотрудничество венгерских специалистов в области информатики.

Организация научно-информационной деятельности в ВНР показана на рис. 124. Государственному комитету непосредственно

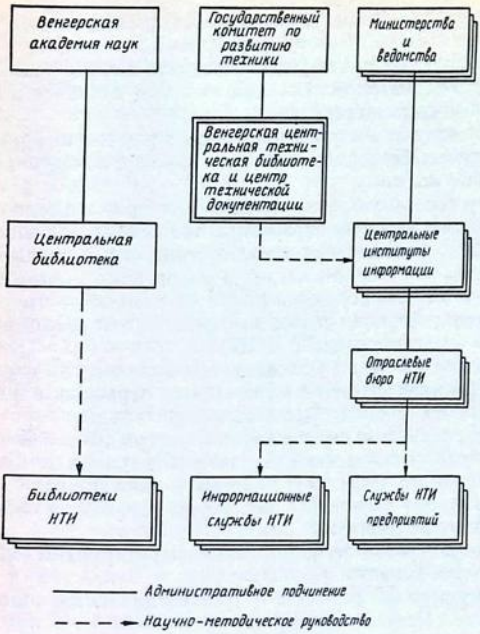



Рис. 124. Организация научно-информационной деятельности в Венгерской Народной Республике

подчиняется Венгерская центральная техническая библиотека с Центром технической документации (ВЦТБ и ЦТД).

В своей работе ВЦТБ выполняет следующие функции:

1) получает и обрабатывает научно-техническую литературу различных областей науки и техники, публикует на этой основе информационные издания;

2) осуществляет различные виды информационного обслуживания государственных учреждений, научно-исследовательских институтов, промышленных предприятий и частных лиц;

3) обобщает отечественный и зарубежный опыт по вопросам организации информационной службы;

4) проводит научно-исследовательскую работу в области научно-технической информации;

5) осуществляет методическое руководство информационными органами;



481-6

6) организует обучение и повышение квалификации работников научно-информационных учреждений;

7) проводит учет и регистрацию переводов научно-технической литературы на венгерский язык, выполненных в любых информационных органах страны;

8) осуществляет международные связи с научно-информационными учреждениями различных стран, а также с международными организациями.

Центральными органами отраслевой информационной сети ВНР являются институты информации или ведомственные органы информации, созданные при министерствах. Они осуществляют руководство всей сетью отраслевых органов информации в своих министерствах. При осуществлении своей деятельности отраслевые органы информации руководствуются методическими разработками и указаниями ВЦТБ и ЦТД.

Следующую ступень в организационной структуре отраслевой сети информационных органов составляют отраслевые информационные бюро. В задачу отраслевых бюро входит:

1) сбор и хранение технической литературы по своей области;

2) обработка поступающей технической периодики и издание информационных документов в форме реферативных листков, обзоров научно-технической литературы и библиографических аннотированных списков;

3) популяризация передового опыта предприятий своей отрасли промышленности;

4) осуществление переводов и организация справочно-информационного обслуживания;

5) контроль за деятельностью органов информации на предприятиях и оказание помощи в получении необходимых информационных изданий.

Нижнюю ступень структуры информационных органов составляют отделы информации или службы информации промышленных предприятий, научно-исследовательских институтов и государственных учреждений, ставящие перед собой задачи:

1) содействовать выполнению плана предприятия, привлекая все возможные средства информационного обслуживания;

2) приобретать и распространять всю техническую литературу, соответствующую профилю предприятия;

3) обобщать и распространять производственный опыт предприятия, направляя информацию об этом в вышестоящий информационный орган;

4) принимать участие во всех изданиях (каталоги, фирменные проспекты, спецификации) предприятия.

Рассмотренная сеть отраслевых информационных органов, возглавляемая Государственным комитетом по развитию техники, тесно сотрудничает с сетью органов научной информации Венгерской Академии наук.

## *Германская Демократическая Республика*

Система информации и документации Германской Демократической Республики (ГДР) в области естественных наук, техники и экономики включает Центральный институт информации и документации, центры информации и документации центральных органов государственного аппарата и центральных научных и научно-технических учреждений, центральные бюро информации и документации в объединениях народных предприятий и отделы информации промышленных предприятий и институтов. Организация научно-информационной деятельности в ГДР показана на рис. 125.

Ответственным за руководство всей информационной и документационной деятельностью в области *естественных наук, техники* и *экономики* является Государственный секретариат по науке и технике. Ему подчинен в качестве руководящего, координирующего и контролирующего центра для названных областей знаний Центральный институт информации и документации (ЦИИД), основанный в 1963 г. Основы этой системы были заложены в постановлении Совета Министров ГДР «О дальнейшем развитии существующей системы информации» от 8. VIII 1963 г.

Принципы построения системы информации и документации в области *общественных наук* были определены постановлением Совета Министров «О дальнейшем развитии существующей в ГДР системы информации и документации в области общественных наук» от 22. IV 1965 г. Для руководства деятельностью в этой области при Немецкой Академии наук был создан Центр информации и документации по общественным наукам.

В задачи ЦИИД входит организация в ГДР единой эффективной системы информации и документации по всем отраслям естественных наук, техники и экономики и планирование информационной и документационной деятельности в этой области. ЦИИД несет ответственность за разработку научных, технических и методических основ информатики и документации и за определение главных направлений подготовки и повышения квалификации информационных кадров. Институт должен обеспечивать координацию информационной и документационной деятельности со стороны центральных государственных органов и отраслевых информационных центров. В функции Института входит также: обеспечение постоянной рационализации информационной деятельности, включая ее механизацию и автоматизацию, координация научно-исследовательской деятельности в области информации и документации и самостоятельное ведение научно-исследовательских работ. Важнейшей задачей Института является проведение анализа современного состояния и путей развития информации в международном масштабе, а также разработка перспектив развития информационной деятельности в ГДР.

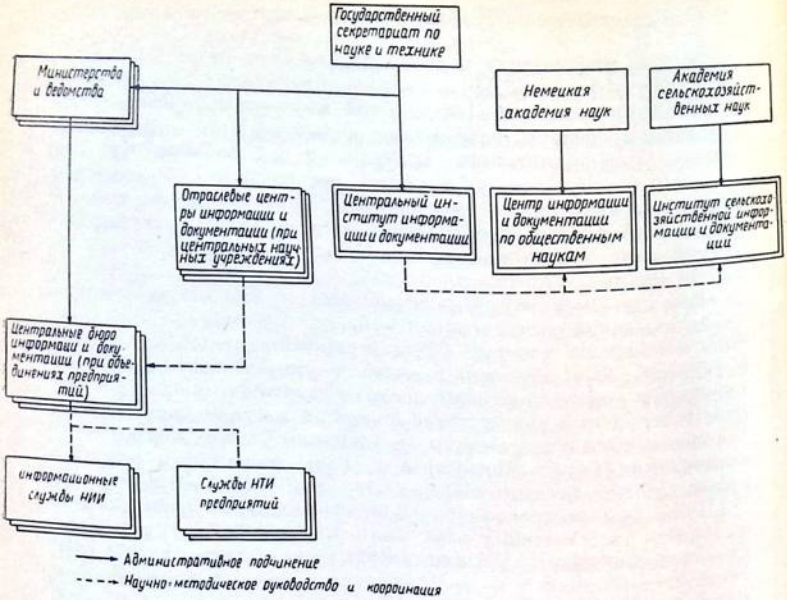



Рис. 125. Организация научно-информационной деятельности в Германской Демократической Республике

В области международного сотрудничества ЦИИД в рамках Совета Экономической Взаимопомощи (СЭВ) активно участвует в работе Постоянной рабочей группы СЭВ по проблемам научно-технической информации, помогает существующим центральным органам государственного аппарата при координировании их международного сотрудничества в области информации и документации, поддерживает многосторонние и двусторонние связи с центральными информационными органами стран — членов СЭВ и двусторонние связи с информационными учреждениями других стран, а также сотрудничает с международными организациями в области информации и документации.

За осуществление информационной и документационной деятельности в центральных органах государственного аппарата, в центральных научных учреждениях, в объединениях народных предприятий и в народных предприятиях несут ответственность соответствующие руководители. Для выполнения соответствующих информационных задач были созданы отраслевые центры информации и документации — на уровне центральных государ-

ственных органов — и центральные бюро информации и документации — на уровне объединений народных предприятий.

Каждый отраслевой центр информации и документации несет ответственность за руководство, координацию и контроль информационной деятельности в какой-либо отрасли народного хозяйства. В их задачи входит: руководство центральными бюро в области организационных и методических вопросов, требующих решений со стороны центрального органа; установление областей, в которых центральные бюро должны осуществлять обработку информационных источников; организация дальнейшего обучения информационных кадров; проведение курсов повышения квалификации и сотрудничество с отраслевыми центрами информации и документации других отраслей народного хозяйства. Отраслевой центр является одновременно информационным органом для своего центрального учреждения.

Каждое центральное бюро информации и документации несет ответственность за руководство, координацию и контроль информационной деятельности в соответствующей узкой отрасли промышленности. В их задачи входит: координация деятельности органов информации предприятий и научно-исследовательских институтов; организация получения, отбора и обработки информационных источников; редактирование и издание реферативных выпусков и других видов публикаций; проведение курсов повышения квалификации и сотрудничество с информационными органами других смежных отраслей, например, предприятиями внешторга, научно-исследовательскими институтами и соответствующими органами информации других отраслей промышленности. Центральные бюро информации и документации одновременно являются информационными органами для объединения народных предприятий.

Основной задачей отдела информации промышленного предприятия или института является обеспечение информацией о научном, техническом и экономическом прогрессе руководителей и всех сотрудников своего предприятия. Отделы информации помогают центральным бюро при обработке соответствующих источников информации и разрабатывают информацию для руководящих кадров.

Для успешного выполнения всех перечисленных задач информационные органы, как правило, имеют собственную техническую библиотеку, группу переводчиков и размножающие устройства.

Реферирование литературы по техническим и естественным наукам, издание реферативных карточек и реферативных журналов в отличие от установившейся практики в других странах — членах СЭВ осуществляется не центральным информационным органом, т. е. Центральным институтом информации и документации, а в основном отделами информации и центральными бюро информации и документации, в отдельных случаях — также

отраслевыми центрами информации и документации. Для реферирования первоисточников привлекаются инженеры или научные работники предприятия или института. Реферативные карточки и реферативные журналы, издаваемые периодически и распространяемые по подписке, были и остаются важнейшим информационным материалом. Однако все чаще наблюдается практика удовлетворения особого спроса потребителей или групп потребителей путем распространения «целенаправленной» информации.

На предприятиях, где нет штатных работников для осуществления информационной деятельности, используются уполномоченные по информации, осуществляющие информационные задачи наряду с их основной производственной деятельностью.

В заключение следует сказать, что информационная система в ГДР построена в соответствии с существующей структурой производства и другими областями техники и экономики. На основании этого отраслевые центры информации центральных государственных органов, в особенности в области промышленности, организуют тесное сотрудничество с информационными учреждениями сферы высшей и специальной школы и научных академий. Подобная информационная система отвечает существующим требованиям народного хозяйства ГДР.

### *Польская Народная Республика*

Общепольская сеть научно-технической и экономической информации состоит из трех ступеней:

1) ведомство,
2) отрасль,
3) предприятие.

Организационная схема сети центров научно-технической информации в Польской Народной Республике (ПНР) представлена на рис. 126. Как видно из приведенной схемы, общее руководство научно-информационной деятельностью в стране возглавляет Комитет науки и техники. Вся сеть информационных звеньев разделена на отдельные группы.

1. Группа центров научно-технической и экономической информации, руководство и контроль за работой которой осуществляет Центральный институт научно-технической и экономической информации (ЦИНТиЭИ). В эту группу входят заводские, отраслевые, ведомственные и воеводские центры технической и экономической информации.

2. Группа центров научной информации Польской академии наук (ПАН), руководство которой осуществляется центром документации и научной информации ПАН.

3. Группа центральных библиотек высших учебных заведений, руководимых Министерством высшего образования.

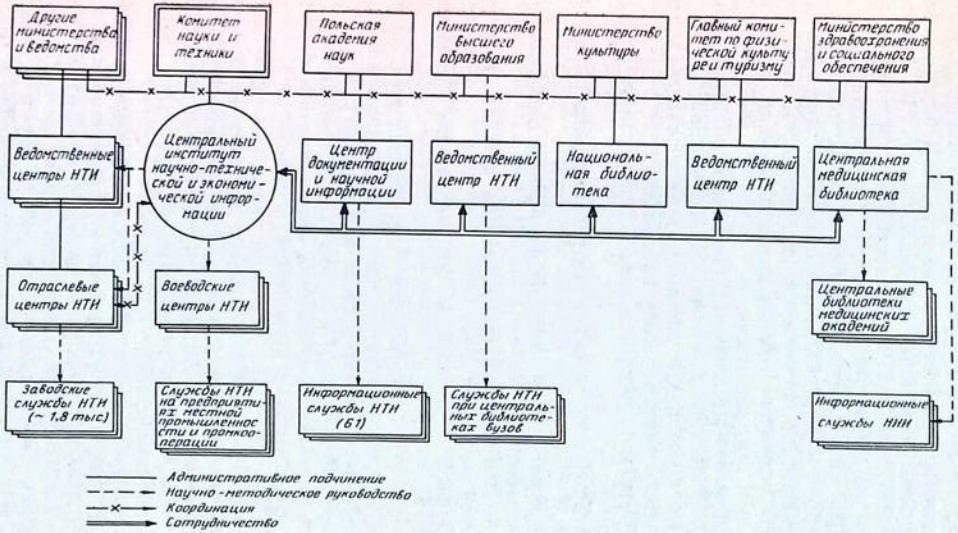



Рис. 126. Организация научно-информационной деятельности в Польской Народной Республике

4. Группа центров научной информации в области медицины, руководство которыми осуществляется Центральной медицинской библиотекой.

Между перечисленными группами информационных центров ПНР существует разделение функций и задач и осуществляется государственная координация их деятельности. В настоящее время в ПНР действует 23 ведомственных центра, 159 отраслевых, примерно 1800 заводских центров и 3 воеводских.

ЦИНТиЭИ был преобразован в 1961 г. из Центрального института научно-технической документации. Он является центральным органом научно-технической и экономической информации в Польше, подчиненным Комитету науки и техники. ЦИНТиЭИ осуществляет следующие задачи:

1) направляет и координирует деятельность польской службы научной, технической и экономической информации;

2) ведет и координирует научные, научно-исследовательские и методические работы в области информатики;

3) координирует информационную деятельность органов всех уровней;

4) руководит комплектованием центральных информационных фондов;

5) выполняет функции центрального органа информации по вопросам информатики;

6) руководит вопросами стандартизации в области научно-технической информации;

7) координирует деятельность по подготовке работников органов информации;

8) координирует сотрудничество информационных органов ПНР с соответствующими зарубежными учреждениями и является их представителем в международных организациях;

9) ведет информационную и издательскую деятельность;

10) сотрудничает с общественными организациями в области развития информационной деятельности в стране.

ЦИНТиЭИ ведет работу в области усовершенствования общепольской системы научно-технической информации и распространения научно разработанных методов информационной работы. В рамках единой системы ЦИНТиЭИ развивает координационную деятельность, стремясь к рациональному разделению труда и исключению параллелизма.

Исследовательские работы ЦИНТиЭИ направлены на решение вопросов теории, методики и практики информационной деятельности. Главной целью этих работ является определение таких методов информатики, при которых содержание, объем и формы информатики способствовали бы прогрессу науки и техники. Методы информатики используются при планировании и реализации научных работ, а также при планировании и управлении народным хозяйством на всех уровнях. Ведется исследование

запросов потребителей информатики с целью выбора соответствующих методов и форм информации. Эти работы осуществляют не только ЦИНТиЭИ, но и некоторые отраслевые центры информации. ЦИНТиЭИ разрабатывает также терминологические и классификационные вопросы.

Соответственно со своими задачами Институт подготавливает и издает информационные обзоры и бюллетени. Кроме традиционной информации о научно-технической литературе, ведется информация о патентах, фирменной литературе, о производстве уникальной аппаратуры, о проектных и опытно-конструкторских работах, о законченных исследовательских работах, отчетах о командировках, о научных конгрессах и съездах, о переводах с иностранных языков и т. д.

В последнее время Институт начал работу по выпуску сигнальной информации о важнейших мировых достижениях науки, техники и производства. Издается серия обзоров в области экономики промышленности, по экономической конъюнктуре, а также о тенденциях и направлениях развития отдельных отраслей народного хозяйства.

Большое место в деятельности ЦИНТиЭИ занимает вопрос подготовки и повышения квалификации работников центров информации. В течение многих лет в Институте работают краткосрочные курсы, осуществляется заочное образование. В области подготовки кадров с ЦИНТиЭИ сотрудничают отраслевые центры информации. По предложению ЦИНТиЭИ в будущем планируется введение в технических, сельскохозяйственных и экономических высших учебных заведениях курса лекций для студентов как будущих потребителей информации.

Органом ЦИНТиЭИ является двухмесячник «Актуальные проблемы информации и документации». Для зарубежных специалистов Институт издает ежеквартальник на русском и английском языках «Обзор польской технической литературы», в котором за год реферируется примерно 800 избранных статей и книг.

*Отраслевые центры* выполняют задания общепольских органов информации по своему профилю, комплектуют необходимые источники информации, обрабатывают их и ведут информационное обслуживание всех заинтересованных учреждений в масштабе страны. Основной формой деятельности отраслевых центров является подготовка и выпуск соответствующих информационных изданий. Кроме того, отраслевые центры составляют тематические подборки материалов, переводят публикации с иностранных языков, следят за текущей литературой по определенным темам, дают консультации и индивидуальные советы, организуют лекции, выставки и совещания для специалистов данной отрасли, а также изготовляют по заказам фотокопии и микрофильмы. Кроме перечисленных заданий, отраслевые центры занимаются

профессиональным контролем обслуживаемых ими заводских центров.

*Заводские центры* выполняют две основные задачи. Прежде всего они обеспечивают специалистов своего предприятия необходимой им технической и экономической информацией. Для этого они приспосабливают к специфическим нуждам своих потребителей материалы, издаваемые отраслевыми центрами. Кроме того, заводские центры систематически информируют соответствующие отраслевые центры о достижениях своих предприятий и их информационных потребностях. Информационные центры с такими функциями создаются не только на производственных предприятиях и в конструкторских и проектных бюро, но и в научно-исследовательских учреждениях.

*Воеводские центры* выполняют следующие задачи:

1) информационная деятельность для президиумов народных советов и предприятий местного хозяйства и промкооперации (на основе материалов отраслевых центров);

2) информационное обслуживание отраслевых центров о достижениях местной промышленности в области технического, экономического и организационного прогресса;

3) организация информационных ячеек в местной промышленности и промкооперации.

Группа центров документации и научной информации *Польской Академии наук* (ПАН) состоит из Центра документации и научной информации ПАН и 61 службы документации и научной информации, действующей при отдельных научных учреждениях ПАН. В задачу Центра документации и научной информации ПАН входит организация информационной деятельности в учреждениях ПАН, осуществление инструктажа и подготовки работников информации ПАН, а также контроль за эффективным функционированием этих центров. Этот центр развивает собственную информационную деятельность, охватывающую организацию науки в стране и за рубежом, проводит научно-исследовательские работы по информатике, издательскому делу; в его ведении находится центральная фототехническая лаборатория, обслуживающая учреждения ПАН. Центры документации и научной информации учреждений ПАН выполняют следующие функции:

1) обработка научных документов своих учреждений;

2) комплектование и обработка зарубежных научных изданий по профилю своего учреждения;

3) предоставление ретроспективной информации на основе справочного аппарата своим фондам;

4) издание библиографических и информационных материалов;

5) изготовление копий документов по заказам сотрудников своего учреждения.

Информационная деятельность в учреждениях *Министерства высшего образования* осуществляется центральными библиотеками

высших учебных заведений (университетов, политехнических институтов, высших сельскохозяйственных и экономических школ). Эта деятельность охватывает, прежде всего, оповещение о новой литературе по профилю вуза и информацию о научно-исследовательских работах, ведущихся в данном вузе. Для этого библиотеки издают информационные списки, составляют библиографические указатели на основе имеющихся каталогов и карточек и по желанию потребителей изготовляют копии документов.

Информационная деятельность в системе учреждений *Министерства здравоохранения и социального обеспечения* осуществляется Центральной медицинской библиотекой и центрами информации научно-исследовательских учреждений (медицинских академий, научно-исследовательских институтов, курсов усовершенствования врачей). Деятельность Главной медицинской библиотеки направлена на составление списков зарубежной медицинской литературы, подготовку и публикацию библиографий польской медицинской литературы, заказ и получение из-за рубежа фотокопий и микрофильмов документов, которых нет в стране, выполнение по заказу фотокопий и микрофильмов имеющихся в библиотеке документов, пропаганду за рубежом важнейших польских работ в области медицины, а также координацию приобретения зарубежных книг и журналов для всех библиотек Министерства здравоохранения и социального обеспечения.

### *Социалистическая Республика Румыния*

Национальная система научной информации Социалистической Республики Румынии (СРР) построена на принципе централизованной государственной координации, централизованном осуществлении научно-информационной деятельности общего и межотраслевого характера и организации информационной деятельности в отдельных отраслях промышленности. Следующая ступень этой системы — органы научной информации в научных институтах и на промышленных предприятиях. Организация научно-информационной деятельности в СРР показана на рис. 127.

Крупные специальные библиотеки рассматриваются как основные звенья общей информационной системы, которые оказывают помощь централизованной обработкой библиографической информации отечественных и зарубежных информационных материалов.

Центральными органами научной информации, охватывающими ряд научных дисциплин или отраслей, являются: Центр научной документации (ЦНД) Академии наук СРР и Центральный институт технической документации (ИТД), подчиненный Управлению метрологии, стандартов и изобретений при Совете Министров СРР. Оба учреждения развертывают комплексную деятельность в области информации, уделяя особое внимание

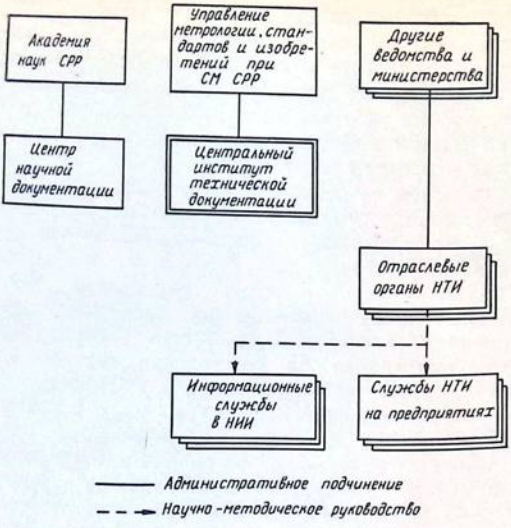



Рис. 127. Организация научно-информационной деятельности в Социалистической Республике Румынии

общей текущей информации о мировой литературе при помощи реферативных журналов и сигнальной информации: ЦНД — в области общественных и естественных наук, ИТД — в области технических наук. ЦНД и ИТД подготавливают и издают реферативные издания о достижениях румынской науки и техники на иностранных языках для зарубежных специалистов. На ЦНД и ИТД возложены разработка и развитие новых методов информационной деятельности, методологическое руководство и повышение квалификации кадров.

Отраслевые органы научной информации СРР охватывают деятельность отдельных министерств или других центральных государственных и кооперативных учреждений. В организационном плане это или отраслевые центры, или бюро. Основная деятельность отраслевых органов информации направлена на создание информационных фондов, обеспечивающих нужды определенной отрасли, подготовку обзоров библиографии, переводов, отражающих интересы данной отрасли, тенденции ее развития и освещение достижений в отечественной и зарубежной практике, обобщение передового производственного опыта, руководство работой органов научной информации в подведомственных учреждениях данной отрасли. Отраслевые центры совместно с центральными органами информации участвуют в подготовке

дров и совершенствовании методов информационной деятельности. В настоящее время в СРР существуют 10 отраслевых центров научной информации.

В научно-исследовательских институтах и конструкторских бюро, а также на большинстве промышленных предприятий существуют отделы или бюро научной информации. Основной задачей этих органов является обеспечение специальной информации по конкретной теме научного исследования или производства. С этой целью они создают фонды информационных материалов в соответствии с нуждами учреждения и координируют деятельность специальных, технических библиотек согласно научно-исследовательским или производственным планам.

### *Чехословацкая Социалистическая Республика*

Общегосударственная система научной, технической и экономической информации Чехословацкой Социалистической Республики (ЧССР) включает информационные центры следующих типов: низовые, специализированные и отраслевые. Организация научно-информационной деятельности в ЧССР показана на рис. 128.

Низовые информационные центры организуются на фабриках, заводах, в учреждениях и т. п. Они обеспечивают всестороннюю информационную деятельность в объеме задач организаций, при которых эти центры созданы. Низовые центры используют в своей работе документационные, исследовательские, методические и другие материалы специализированных и отраслевых центров. В целях распространения информации о новых достижениях науки и техники они методически направляют деятельность технических информаторов — широкого актива сотрудников в отдельных подразделениях завода, фабрики или учреждения. Органы общественных организаций предприятия принимают участие в разработке центрами проекта совместного плана научной и производственно-технической пропаганды, утверждаемого руководителем и соответствующим органом общественной организации.

Специализированные информационные центры организуются, как правило, при ведущих специализированных центрах развития науки и техники.

Отраслевые информационные центры организуются не только при научно-исследовательских институтах или других крупных центрах, но и как самостоятельные институты информации по общеотраслевой или межотраслевой проблематике.

Организация специализированных и отраслевых центров и определение информационных задач в отраслях для отдельных центров утверждается по представлению соответствующих центральных органов и Центра научной, технической и экономической

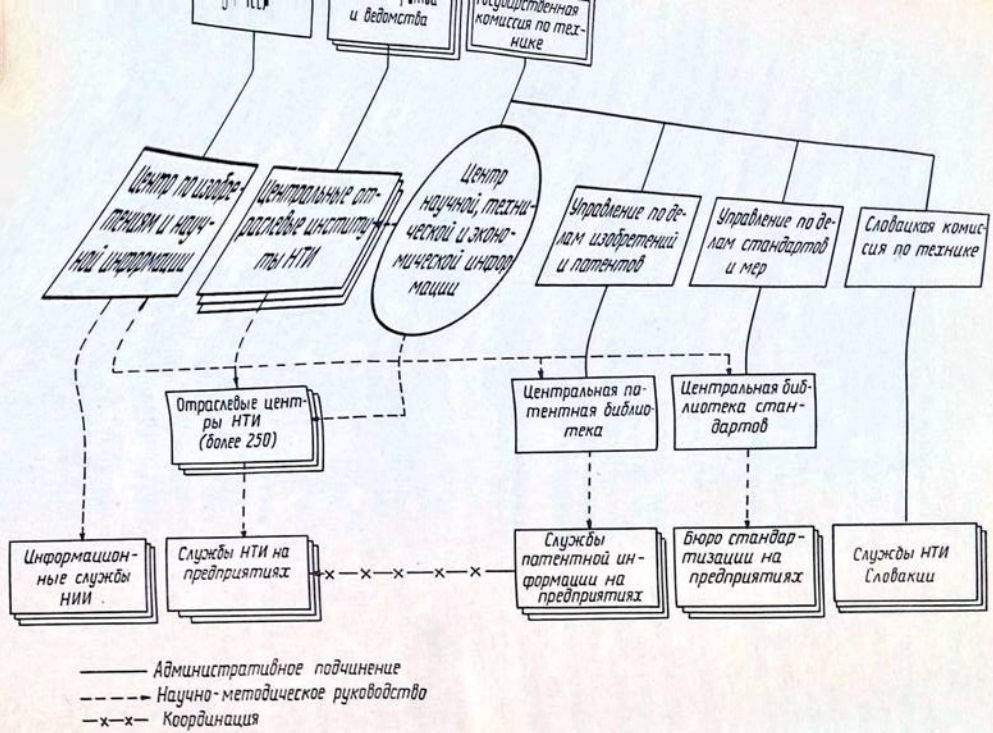



Рис. 128. Организация научно-информационной деятельности в Чехословацкой Социалистической Республике

информации Государственной комиссией по технике. Установленные разграничения в информационной деятельности по отдельным отраслям уточняются путем договоренности между отдельными специализированными и отраслевыми информационными центрами.

Основное содержание деятельности информационных центров и взаимные отношения между ними определяются их статутами. Статуты отраслевых центров утверждаются соответствующими центральными органами по предварительному согласованию с Государственной комиссией по технике, которое осуществляется по представлению Центра научной, технической и экономической информации. Общегосударственная научно-техническая информационная система ЧССР в масштабах всей страны управляется централизованно Государственной комиссией по технике, которая является центральным органом управления информацией.

Государственная комиссия также координирует и ориентирует деятельность других элементов центрального управления информационной деятельностью — Словацкой комиссии по технике, Управления патентов и изобретений, Управления стандартизации и мер. Словацкая комиссия по технике в качестве органа Словацкого Национального Совета и областного органа Государственной комиссии наблюдает за выполнением задач и развитием информационной деятельности в Словакии, обеспечивая единство организации и управления в рамках общегосударственной информационной системы.

Центр научной, технической и экономической информации направляет и контролирует деятельность в области научной, технической и экономической информации, обеспечивает информационную деятельность, которая должна осуществляться централизованно, осуществляет исследования в области научной, технической и экономической информации и ориентирует научно-техническую пропаганду.

Кирпичева И. К. Библиография в помощь научной работе. Методическое и справочное пособие. Л., 1958.
Коновалова М. Н. Алфавитно-предметный указатель к систематическому каталогу. Пособие для библиотекарей. Л., 1955.
Кричевский Г. Г. Общие библиографии зарубежных стран. М., 1962.
Немировский Е. Л. Патентная литература. Учебное пособие. М., ЦНИИПИ, 1964.
Симон К. Р. История иностранной библиографии. М., 1963.
Шамурин Е. И. Словарь книговедческих терминов. Для библиотекарей, библиографов, работников печати и книжной торговли. М., 1958.
Barker R. E. Le livre dans le monde. Etude sur le commerce intrenational du livre. Paris, Unesco, 1957.
Brown C. H. Scientific serials. Chicago, ACRL, 1956.
Collison R. L. Encyclopaedias, their history throughout the ages. A bibliographical guide. New York, 1962.
A guide to the world's abstracting and indexing services in science and technology. Washington, 1963.
Malclès L.-N. La bibliographie. Paris, 1956.

*К главе 3*

Брискман М. А. и Бронштейн М. П. Составление библиографических пособий. Практическое руководство. М., 1964.
Герман П. Практическое применение Универсальной десятичной классификации. Техника классификации. Пер. с нем., изд. 4, доп. и переработ. М., 1964.
Гиляревский Р. С. Библиографическое описание как элемент информации. М., 1961. (ВИНИТИ. Доклады на Конференции по обработке информации, машинному переводу и автоматическому чтению текста, вып. 6).
Гиляревский Р. С. Развитие современных принципов книгоописания. Краткий очерк. М., 1961.
Гиляревский Р. С. Отражение литературы на иностранных языках в библиотечных каталогах. Методические материалы. М., 1963.
Гиляревский Р. С. Описание произведений печати. Алфавитный каталог. Методическое пособие. М., 1965.
Гиляревский Р. С. и Гривнин В. С. Определитель языков мира по письменностям. Изд. 3, испр. и доп. М., «Наука», 1965.
Гиляревский Р. С. и Старостин Б. А. Иностранные имена и названия в русском тексте. Справочник. М., «Международные отношения», 1968.
Горькова В. И. и Муранивский Т. В. Пособие по реферированию и редактированию научно-технической литературы. (Справочник для редакторов, референтов и корректоров). М., 1964.
Заполъская Н. Н. Классификация произведений печати и организации систематического каталога в технических библиотеках. Методическое пособие. М., 1964.
Коллисон Р. Неисчезающий барьер. Переводы как средство общения между Востоком и Западом: анализ указателя переводов Index translationum. «Бюлл. ЮНЕСКО для библиотек», 1962, т. 16, № 6.
Кругликова В. П. Предметизация произведений печати. Общая методика. М., «Книга», 1967.
Марченко Э. В. Общая методика предметизации произведений печати. Лекция для студентов. М., 1956.
Мельчук И. А. и Равич Р. Д. Автоматический перевод 1949—1963. Критико-библиографический справочник. М., ВИНИТИ, 1967.
Михайлов А. И. и Тареев Б. М. Научная информация в области электротехники и энергетики. Изд. 2, переработ. М., 1964.

Немченко В. В. и Сытина Н. Я. Описание специальных видов технической литературы. Методическое и практическое пособие. М., 1963.
Никифоровская Н. А. (сост.). Правила библиографического описания произведений печати. Пособие для библиографов, научных и издательских работников. Л., 1964.
Новикова Е. А. Основы книгоописания и организации алфавитного каталога. М., 1959.
Новикова Е. А. Описание произведений печати и организация алфавитного каталога в технических библиотеках. Методическое пособие. М., 1964.
Руководство к реферированию научной литературы по ботанике, физиологии растений, растениеводству, лесоводству, почвоведению и агрохимии. Под ред. Л. Л. Балашова. М., ВИНИТИ, 1964.
Садова М. А. В поисках книги. Библиографические разыскания при выполнении читательских требований. М., 1963.
Соловьев В. И. Редактирование реферата. Учебное пособие. Под общ. ред. Н. М. Сикорского. М., МПИ, 1965.
Успенский В. А. К проблеме транслитерации русских текстов латинскими буквами. «Научно-техн. информация», 1967, серия 2, № 7.
Фирсов Г. Г. Регистрация и описание периодических изданий. М., «Сов. Россия», 1961.
Шамурин Е. И. Методика составления аннотаций. М., 1959.
Шамурин Е. И. Очерки по истории библиотечно-библиографической классификации, т. 1—2. М., 1955—1959.
Ashworth W. (Gen. ed.) Handbook of special librarianship and information work. 2nd ed. London, Aslib, 1962.
Grolier E. de. A study of general categories applicable to classification and coding in documentation. Paris, Unesco, 1962.
Metcalfe J. Subject classification and indexing of libraries and literature. New York, Scarecrow Press, 1959.
Vickery B. C. Classification and indexing in science. 2nd ed. London, Butterworths, 1959.

*К главам 4, 5, 6*

Арапов М. В. Применение метода тезаурусов к машинному переводу (Обзор зарубежной литературы). «Научно-техн. информация», 1962, № 11.
Арапов М. В. Некоторые принципы построения словаря типа «тезаурус». «Научно-техн. информация», 1964, № 4.
Бернштейн Э. С. Формализованный язык и критерий смыслового соответствия информационно-поисковой системы «Пусто-Непусто-4». «Научно-техн. информация», 1963, № 12.
Бернштейн Э. С., Лахути Д. Г., Чернявский В. С. О критериях оценки поисковых систем. «Научно-техн. информация», 1964, № 3.
Бернштейн Э. С., Лахути Д. Г. и Чернявский В. С. Вопросы теории поисковых систем. Под общ. ред. В. Чернявского. М., ВНИИЭМ, 1966.
Влэдуц Г. Э. О некоторых сторонах исследований по созданию информационно-поисковых систем. «Научно-техн. информация», 1961, № 1.
Всесоюзная конференция по информационно-поисковым системам и автоматизированной обработке научно-технической информации, 3-я, т. 1—4. Под общ. ред. А. И. Михайлова и др. М., ВИНИТИ, 1967.
Грязнухина Т. А., Пшеничная Л. Э., Скороходько Э. Ф. Система информационного поиска. Киев, «Наукова думка», 1964.

Poindron P. (Ed.). Manuel pratique de reproduction documentaire et de selection. Publ. FID 353. Paris, Gauthier-Villars, 1964.
Punched cards. Their applications to science and industry. R. S. Casey, J. W. Perry, M. M. Berry, A. Kent (Eds.) 2nd. ed. New York, Reinhold, 1959. — См. русск. перевод: Перфорированные карты и их применение в науке и технике. М., Машгиз, 1963.
Scheele M. Punched-card methods in research and documentation (With special reference to biology). New York, Interscience Publishers, 1961.
Shera J. H., Kent A., Perry J. W. (Eds.). Information systems in documentation. New York, Interscience Publishers, 1957.
Singer T. E. R. (Ed.). Information and communication practice in industry. New York, Reinhold Publishing Corp., 1958.

*К главе 8*

Автоматизация в полиграфии. Сборник переводов. М., 1961.
Букатин Е. А. Микрорепродукция и ее применение в службе научно-технической информации. «Научно-техн. информация», 1962, № 4.
Вопросы электрографии. Сборник переводов. М., 1960.
Средства копирования и размножения. М., 1965. (ВИНИТИ. Сборник переводов, № 3).
Гельман-Виноградов К. Б. Микрофотокопирование документальных материалов и организация работы с микрофотокопиями в архивах СССР. М., 1961.
Жилевич И. И., Немировский Е. Л. Электрофотография. М., «Искусство», 1961.
Интероргтехника-66. Средства механизации инженерно-технических и административно-управленческих работ. Международная выставка. Советские экспонаты. Каталог. М., 1966.
Кальмансон В. А. Современная техника набора текста информационных изданий. «Научно-техн. информация», 1963, № 2.
Кальмансон В. А. Быстродействующие печатающие устройства электронных вычислительных машин. М., Информстандартэлектро, 1967.
Левинсон Н. Г. Механизация управленческого труда (оргатехника), т. 1. М., Машгиз, 1958.
Наймарк А. М. Подготовка информационных материалов к изданию методом макет-оригинала. «Научно-техн. информация», 1962, № 6.
Стибиц В. От механизации к автоматизации управленческих работ. М., Машгиз, 1963.
Фридман И. М. Микрофотокопирование. М., «Искусство», 1956.
Шицгал А. Г. Микроиздания. М., ВНИИПИ, 1962.
Bartlett H. A. A guide to reproductive processes. «Office Methods and Machines», 1965, v. 12, № 142.
Compendium of document reproduction. «The Industrial and Commercial Photographer». Suppl., 1963, v. 3, № 4.
Lewis C. M., Offenhauser W. H., jr. Microrecording: Industrial and library application. New York, Interscience Publishers, 1956.
Milward G. E. (Ed.) Organisation and methods. London, Macmillan and Co., Ltd., 1960. — См. русск. перевод: Дж. Милвард (Ред.). Рационализация управления и техника конторской работы. М., ИЛ, 1963.
Poindron P. (Ed.). Manuel pratique de reproduction documentaire et de selection. Publ. FID 353. Paris, Gauthier-Villars, 1964.
The revolution in office copying. Special report. «Chemical and Engineering News», 1964, v. 42. № 28—29.
Verry H. R. Document copying and reproduction processes. London, Fountain Press, 1960.

Verry H. R. Microcopying methods. London—New York, Focal Press, 1963.
Wysocki A. Mała poligrafia. Wrocław, Ossolineum, 1965. (Książki o książce).

*К главе 9*

Арутюнов Н. Б. О мерах по улучшению организации научно-технической информации в стране. «Научно-техн. информация», 1962, № 5.
Арутюнов Н. Б. Совершенствование организации научно-информационной деятельности в стране и задачи ВИНИТИ. «Научно-техн. информация», 1963, № 2.
Арутюнов Н. Б. О мерах по дальнейшему совершенствованию научно-технической информации в стране. «Научно-техн. информация», 1964, № 6.
Арутюнов Н. Б. Дальнейшее развитие системы научно-технической информации. «Научно-техн. информация», 1967, серия 1, № 11.
Михайлов А. И. Десятилетие ВИНИТИ и перспективы его дальнейшего развития. «Научно-техн. информация», 1962, № 12.
Михайлов А. И. Проблемы научной информации. «Научно-техн. информация», 1963, № 1.
Михайлов А. И., Черный А. И. и Гиляревский Р. С. Развитие информатики в СССР. «Научно-техн. информация», 1967, серия 2, № 11.
Остромецкий А. Некоторые вопросы создания и использования справочно-информационных фондов в НИИ, КБ и на предприятиях. Материалы по научно-техн. информации. М., ВИНИТИ, 1966.
Почкай И. Б. и Чашников И. П. В. И. Ленин и научная информация. М., ВИНИТИ, 1965.
Служба научно-технической и экономической информации в Польше. Варшава, CIINTE, 1966.
Теплов Д. Ю. Информационно-техническая библиография в СССР. Краткий исторический очерк. М.—Л., Изд-во АН СССР, 1962.
Томан И. Техническая информация в Чехословакии. (Практика информационной работы). М., ЦИНТИ Электропром, 1963.
Циганик М. Организация и использование фонда электротехнической информации в науке, технике и экономике. (Сокр. перевод со словацкого). М., ВНИИЭМ, 1967.

| | |
|---|---|
| RADIR | Random Access Document Indexing and Retrieval (Система индексирования и поиска документов с произвольным доступом) |
| SLIC | Selective Listing in Combination (Избирательное перечисление комбинаций) |
| TDCK | Technisch Documentatie- en Informatie- Centrum voor Krijgsmacht (Центр технической документации и информации Вооруженных сил Нидерландов) |
| UCS | Universal Card Scanner (Универсальное устройство для сканирования перфокарт) |
| WADEX | Word & Authors inDEX (Указатель ключевых слов и авторов) |

## ИМЕННОЙ УКАЗАТЕЛЬ


Авакян Э. 603
Аврамеску А. 49, 729
Аграев В. А. 231, 233
Азимов А. 149
Акофф Р. 35
Аламбер Ж. д' 399
Александров П. С. 373
Алисов М. И. 654
Аллен Г. 142
Аллен Л. 515
Амбарцумян З. Н. 209, 322, 331, 361, 729
Амера Синха 417
Анастасевич В. Г. 162
Андрьюс Д. 509
Арапов М. В. 421, 731
Аристофан Византийский 416
Арутюнов Н. Б. 735
Атертон П. 52, 95, 261, 271, 342, 431, 469
Ахманова О. С. 257, 263
Ашуэрт У. 731
Ашшурбанипал 69, 160, 322

Багг Т. 590, 591
Байглоу Дж. 17
Байрон Дж. Г. 397
Балашов Л. Л. 731
Балли Ш. 418, 419, 439
Бальц К. 423
Барабанов А. Н. 99
Баркер Р. 85, 730
Бартлетт Г. 734
Бартон Р. 97, 150
Бар-Хиллел И. 39, 64, 286, 302
Баттен У. 382
Бейкер Д. 28, 34
Бейли М. 460
Беккер Дж. 48, 732, 733
Белло Ф. 38, 568, 729
Белпорт Б. 479
Белцер Дж. 307, 308
Берг А. И. 64, 65
Бердичевская Ц. М. 729
Берков П. Н. 729
Бернал Дж. 29, 72, 91, 729
Бернштейн Г. 395
Бернштейн Э. С. 265, 297, 433, 434, 731
Бернье Ч. Л. 346, 420, 446, 447
Берри М. 384, 433, 459, 732, 734
Бестермен Т. 117
Бирюков Б. В. 445
Блисс Г. П. 205—208
Блумфилд Л. 399
Боз М. 732
Болл Г. 590
Боннар А. 251
Борко Г. 310
Бори Ч. 33, 34, 561, 732
Бородин В. В. 231, 233
Боспер Р. 419
Боуден Б. 397
Браге Т. 244
Браун Ч. 85, 730
Браунсон Э. 48
Брейер А. 624
Бреннер С. 384
Бреннер Э. 498
Бриллюэн Л. 55, 62, 63, 729
Брискман М. А. 729, 730
Бродхерст П. 392
Бронштейн М. П. 730
Брукс Р. 515
Брэдфорд С. 47, 93—96
Брэдшоу П. 289, 583
Брэндхорст У. 294
Брюне Ж. Ш. 182, 184
Будагов Р. А. 259
Буерн У. 306
Букатин Е. А. 734
Булаховский Л. А. 399, 410
Бунтрок Г. 459, 503
Бут А. Д. 236
Буш В. 591
Бьоркбом К. 46
Бьюшоп У. 583, 603
Бэббедж Ч. 397

Устройства для реализации информационно-поисковых систем 535—620
Учебники 87
Учетно-регистрационная библиография 115
Файлсёрч, информационно-поисковая система 593—597
Фактографические ИПС 248
Фасетная классификация 201—205, 338—344
Фильморекс, информационно-поисковая система 562—568
Фирменные каталоги 103
Флексорайтер, наборно-пишущая машина 656
ФЛИП, информационно-поисковое устройство 609—611
Фонетика 256
Фотографические процессы 623—625
Фотокопирование 623—625
Фотолист, устройство для изготовления печатных каталогов 124
Фотон, фотонаборная машина 657—658
Фотонабор 657—659
Фототермография 633—634
Хедлайнер, фотонаборная установка 655
Хитуол, информационно-поисковая система 616—617
Цепные алфавитно-предметные указатели 204, 361—365
Центральная междуведомственная комиссия по закупке и распределению заграничной литературы (Комиполит) 667
Центральный институт научной информации по строительству и архитектуре (ЦИНИС) 695—696
Центральный каталог 123
Центральный научно-исследовательский институт патентной информации и технико-экономических исследований (ЦНИИПИ) 689—690
Цитированная литература, указатель 141—153
Чехословакия, организация информационной деятельности 725—727
Читальные аппараты 625
Читательский каталог 123
Щелевые перфокарты 550—551
Экспериментальная информационная машина ВИНИТИ 552—553
Экспериментальный микрофильмовый селектор фирмы ИБМ 611, 614—615
Эктафакс, процесс 636
Экспресс-информация 113—114, 682—683
Электрографические процессы 637—644
Электроискровые копировальные аппараты 660—661
Электролитический процесс 643
Электронные цифровые вычислительные машины
— роль в механизации и автоматизации информационного поиска 396—398
Электрофакс, процесс 641—642
Энциклопедии 108—110
Эффективность документальных ИПС — методы оценки 302—315
Юниверсал Кард Скеннер, информационно-поисковая система 555—558
Язык RX-кодов 460—461

# ОГЛАВЛЕНИЕ


Предисловие к первому изданию . . . . . 5

Предисловие авторов ко второму изданию . . . . . 9

Глава 1. Предмет и метод информатики . . . . . 12

Роль науки в современном обществе . . . . . 12
Закономерности развития науки и научная информация . . . 16
Потребности общественной практики и научная информация . . 31
Информатика как самостоятельная научная дисциплина . . . . 42
Информатика и другие научные дисциплины . . . . . 61

Глава 2. Документальные источники научной информации . . . 81

Научные документы и их виды . . . . . 81
Первичные документы и издания . . . . . 83
Вторичные документы и издания . . . . . 108

Глава 3. Аналитико-синтетическая переработка документов . . . 158

Библиографическое описание . . . . . 159
Классификация (систематизация) . . . . . 179
Предметизация . . . . . 220
Аннотирование . . . . . 222
Реферирование . . . . . 224
Научный перевод . . . . . 235

Глава 4. Информационно-поисковые системы . . . . . 244

Информационный поиск. Общие принципы и понятия . . . . . 244
Информационно-поисковый язык . . . . . 255
Критерий смыслового соответствия . . . . . 282
Методы оценки эффективности документальных информационно-поисковых систем . . . . . 302

Глава 5. Традиционные информационно-поисковые языки . . . . 316

Общие принципы классификации . . . . . 318
Иерархические библиотечно-библиографические классификации 326
Фасетные библиотечно-библиографические классификации . . . 338
Нотация и алфавитные указатели к классификационным схемам 344
Алфавитно-предметные классификации . . . . . 346

Глава 6. Дескрипторные информационно-поисковые языки . . . 367

Общие принципы и понятия . . . . . 367
Синонимия ключевых слов и ее устранение . . . . . 398
Многозначность ключевых слов и ее устранение . . . . . 408
Общие понятия тезауруса . . . . . 413
Парадигматические отношения между дескрипторами . . . . . 432
Способы выражения парадигматических отношений в тезаурусах 455
Описание основных зарубежных тезаурусов . . . . . 469

Обобщенная структурная схема тезауруса . . . . . .
Синтагматические отношения между дескрипторами . . . . . .
Специальные информационно-поисковые языки . . . . . .

Глава 7. Методы и средства реализации информационно-поисковых систем . . . . . .
Общие принципы . . . . . .
Информационно-поисковые устройства с дискретными носителями информации . . . . . .
Информационно-поисковые устройства с непрерывными носителями информации . . . . . .

Глава 8. Копирование и размножение документов (репрография)
Копирование документов . . . . . .
Размножение документов . . . . . .

Глава 9. Организация научно-информационной деятельности . .
Научно-информационная деятельность в СССР . . . . . .
Научно-информационная деятельность в странах—членах СЭВ

Библиография . . . . . .
Указатель сокращений . . . . . .
Именной указатель . . . . . .
Предметный указатель . . . . . .

*Александр Иванович Михайлов, Аркадий Иванович Черный, Руджеро Сергеевич Гиляревский*

**Основы информатики**

Второе, переработанное и дополненное издание

*Утверждено к печати*
*Всесоюзным институтом научной и технической информации Академии наук СССР*

Редактор издательства *А. Н. Знаменская*
Технический редактор *И. Н. Жмуркина*

Сдано в набор 29/I 1968 г. Подписано к печати 22/IV 1968 г. Формат 60×90¹/₁₆. Усл. печ. л. 47,35. Уч.-изд. л. 48,7. Тираж 25000 экз. Т-07506 Тип. зак. 864. *Цена 3 р. 45 к.*
Издательство «Наука». Москва, К-62, Подсосенский пер., 21

1-я типография издательства «Наука». Ленинград, В-34, 9 линия, д. 12