

# МИРОВАЯ ИНДУСТРИЯ ОНЛАЙНОВЫХ БАЗ ДАННЫХ \*

Л. К. БОБРОВ

*Государственная публичная научно-техническая библиотека СО РАН*

*Новосибирск, Россия*

e-mail: bobrov@libr.nsk.su

The survey of the current state of the informational centers and computer-oriented data bases of the "global" character (large multi-field data based, generic data bases such as patent information data bases, special-purpose data bases on science and technology etc.) The problems of the access to such bases via world networks are also discussed.

## 1. Введение

Начиная со второй половины двадцатого столетия мировое развитие начало сопровождаться бурным ростом количества научной и технической информации. Этот феномен получил название "информационного взрыва", и его обсуждению было посвящено большое количество научных работ (см., напр., фундаментальные труды [1–6]). В частности, в отношении характеристик мирового информационного потока отмечалось, что ежегодно в мире публикуются результаты примерно двух миллионов исследований, ежеминутно появляется около двух тысяч страниц печатной информации, которая публикуется на шестидесяти языках [2] в более чем ста тысячах периодических изданий [7]. В связи с этим специалист, стремящийся иметь информацию о всех новых достижениях в своей области, вынужден тратить до трети своего рабочего времени на подбор, анализ и изучение информации, иначе существует опасность непроизводительных затрат до 45 % отпущенных на разработки и исследования средств [8].

Изучая проблему "информационного взрыва", многие ученые пришли к выводу о том, что количество публикуемой информации возрастает экспоненциально, и в соответствии с этим рост информации к концу текущего столетия должен замедлиться. Однако это предположение не оправдалось, и, более того, как полагают авторы работы [4], скорость роста количества научной и технической информации, возможно, даже увеличивается. Вероятно, данную ситуацию можно объяснить тем, что бурное развитие средств вычислительной техники, наблюдаемое в два последних десятилетия, привело к не учитывавшемуся ранее мощному потоку непубликуемой информации, отражающей результаты постоянно расширяющихся естественнонаучных исследований. Например, в Системе спутникового наблюдения за Землей, которое ведет Национальное управление по аэронавтике и освоению космического пространства (НАСА) США, объем ежедневно появляющейся информации

---

\* © Л. К. Бобров, 1997.

эквивалентен более чем 100 млн страниц печатной продукции. К тому же информатизация различных сфер человеческой деятельности способствовала росту потоков финансовой, технико-экономической, технологической, производственной и другой информации, а также породила новые формы использования этой информации [4].

Данные факторы вызвали новый скачок исследований по проблемам, связанным с информационным обеспечением различных сфер человеческой деятельности (наука, экономика, управление, производство, социальная сфера и др.). Информация стала рассматриваться как один из главных видов ресурсов, определяющих экономический статус государства. В соответствии с этим с конца семидесятых годов начались исследования в области управления информационными ресурсами (information resources management или information management). В последние годы вместо терминов "информация" и "информационные ресурсы" все чаще стали употреблять термин "ресурсы знаний" (knowledge resources), а информационные работники стали называться "специалистами по знаниям" (knowledge officers). Ведущие западные специалисты в области экономики и менеджмента считают, что обеспечение успеха фирмы на рынке напрямую зависит от степени овладения ею новой информацией, новыми знаниями. Для этого фирма должна стать "обучающейся" (learning organization) и способной эффективно использовать современные информационные технологии с целью получения новых знаний [4, 9].

Таким образом, информация является в настоящее время ресурсом, играющим доминирующую роль в системе глобального мирового развития. При этом наблюдается постоянное совершенствование систем представления информации и способов ее доведения до потребителей. Неуклонно растет количество разнообразных баз данных, совершенствуются носители информации, развиваются сети передачи данных. По этой причине в дальнейшем мы будем вести речь, в основном ориентируясь на эти бурно развивающиеся направления, хотя, конечно же, процессы данного развития протекают на фоне улучшения и традиционных форм и методов фиксации информации (например, в виде печатных изданий) и ее распространения.

## **2. Структура информационного рынка, области и секторы рынка**

Информация является идеальной, а не материальной категорией, она не производится специально для продажи как материальные продукты труда. Поэтому в строгом смысле информация товаром считаться не может. Сама информация не продается и не покупается, продаются и покупаются содержащие ее материальные носители и право на использование данной информации. Объектом купли-продажи служат также информационные услуги, основанные на использовании различных источников информации. Процессы купли-продажи регламентируются в отличие от системы товарного производства законом об интеллектуальной собственности, законом об авторском праве и другими нормативными документами. Информационную отрасль можно считать инфраструктурной отраслью, выполняющей обеспечивающие функции (как, например, транспорт и связь). Однако, в отличие от других отраслей, информационный рынок является в настоящее время одним из наиболее динамичных, а его экономические аспекты остаются во многом неясными и требуют дальнейших углубленных исследований [4].

Структура современного информационного рынка в большей степени отражает рынок информационных услуг. Поэтому перечень приведенных ниже областей и секторов современного информационного рынка не является исчерпывающим. На практике данный ры-

нок включает или частично охватывает и другие области, секторы и сегменты. Структура и наполнение рынка постоянно совершенствуются, развиваются и пополняются.

Область информационной деятельности охватывает секторы деловой, социально-политической, научно-профессиональной, массовой и потребительской информации.

Сектор деловой информации включает биржевую и финансовую информацию (котировки ценных бумаг, информацию о рынке капиталов, товаров, инвестиций, курсы валют, цены, учетные ставки и т.п.), экономическую и статистическую информацию (числовые, экономические, социальные, демографические и другие показатели), коммерческую информацию (данные о фирмах, компаниях, корпорациях, товарах и т.п.) и деловые новости в области экономики, бизнеса и менеджмента [4, 5].

Сектор социально-политической информации ориентирован на обслуживание органов власти и управления информацией статистического, демографического, социального характера и др.

Сектор научно-профессиональной информации охватывает все виды информации в сфере науки, производства и общественных (социальных) отношений.

Сектор массовой и потребительской информации включает сведения массового характера (развлекательная информация, информация о погоде, расписаниях движения транспорта, справочники-путеводители, и пр.), а также справочники, энциклопедии, условия размещения вкладов в различных банках, данные о ценах на товары и услуги, телетекст, видеотекст и др.

Область электронных сделок включает системы электронных банковских операций (базирующихся на уже проверенных практикой стандартах и отработанных технических решениях), электронные биржи и системы заказа и резервирования продукции и услуг — например, системы заказа и бронирования авиа- и железнодорожных билетов [5].

Область вычислительной техники и электронных коммуникаций охватывает всевозможные технические средства (начиная от информационно-вычислительных комплексов, рабочих станций и профессионально ориентированных автоматизированных рабочих мест и до различного рода сетевого оборудования — модемов, адаптеров, карт и т. п.), программные и информационные (информационно-поисковые) системы, системы управления базами и банками данных, средства поддержки сетевых режимов работы, а также собственно сети передачи данных и технологии решения в этих сетях различных задач (например, реализация систем телеконференций, электронной почты, и т. п.).

### **3. Организация информационной деятельности и объекты информационного рынка**

Как уже отмечалось, информационная отрасль является отраслью инфраструктурной и чрезвычайно важной, что определяет то серьезное внимание, которое уделяется данной отрасли государственными органами. В большинстве развитых стран наряду с большим количеством относительно маломощных и полностью самостоятельно функционирующих информационных фирм можно выделить три уровня информационных органов, работа которых проводится под эгидой (в той или иной степени, формах и проявлениях) государства [5].

Первый уровень составляют национальные информационные центры, готовящие машиночитаемые базы данных "глобального" характера (крупные политематические базы данных, видовые базы данных, например базы патентной информации, тематические базы

данных по науке и технике и т. п.) и (или) предоставляющие доступ к базам данных через мировые сети передачи данных.

Второй уровень образуют отраслевые и территориальные органы информации, готовые соответственно отраслевые либо территориальные базы данных и обслуживающие свои отрасли (территории) как путем предоставления баз данных собственной генерации, так и используя базы данных органов первого уровня.

Третий уровень составляют информационные брокеры, обслуживающие конечных потребителей путем использования баз данных, производимых центрами первого и второго уровней.

Следует отметить, что большинство независимых коммерческих информационных организаций второго и третьего уровней работают в области деловой, массовой и потребительской информации, поскольку именно эти области дают возможность безубыточной деятельности.

Основным товаром информационного рынка, как, впрочем, и основой мировой информационной индустрии, являются базы данных, производимые так называемыми центрами-генераторами баз данных и доводимые до потребителей через центры обработки (центры-распределители) баз данных, шлюзовые службы доступа к базам данных и разветвленную сеть брокеров и других структур аналогичной направленности.

Центры-генераторы баз данных (database publishers) основное внимание уделяют отбору материалов, отражаемых в базах данных, и качеству наполнения документов (записей). При этом центр-генератор готовит информацию в машиночитаемом виде в соответствии с существующими стандартами (регламентирующими физическую структуру и, например в части библиографической информации, — наполнение записей, то есть состав и содержание полей) и далеко не всегда занимается формированием баз данных. Эта процедура зачастую осуществляется специализированными организациями — так называемыми центрами обработки баз данных — по контрактам с центрами-генераторами.

Центры обработки баз данных, иначе называемые вендорами (vendors) или хост-центрами (hosts), осуществляют обработку данных, получаемых от центров-генераторов, с целью создания баз данных, наиболее эффективных с точки зрения поисковых характеристик и органично вписывающихся в используемую вендором программную среду. Крупные вендоры поддерживают одновременно сотни различных баз данных и предоставляют эти базы в теледоступ для большого числа пользователей. В последнее десятилетие прочное место на информационном рынке заняли вендоры, специализирующиеся на производстве баз данных на компактных оптических дисках (CD-ROM) и реализующие их через относительно небольшое число дилерских фирм, охватывающих своей деятельностью все регионы мира. В этот же временной период на информационном рынке появились коммерческие службы — так называемые шлюзовые службы доступа [5], предлагающие пользователям сетевое подключение к тысячам различных баз данных через упрощенный меню-ориентированный интерфейс. Заметим, что нередки случаи (особенно в России), когда центры-генераторы самостоятельно формируют свои базы данных и организуют доступ к ним.

Поиск информации, соответствующей информационной потребности клиента, — операция нетривиальная. Большинство пользователей не имеют необходимого опыта, навыков и знаний для качественного проведения поиска, хотя, на первый взгляд, наличие дружественного к пользователю интерфейса в большинстве информационных систем создает обратное впечатление. На самом же деле, проводя поиск самостоятельно, пользователь либо получает большое количество "шумовой", не нужной ему информации, либо извле-

кает из баз данных лишь небольшую часть действительно необходимых ему сведений. Поэтому присутствие на информационном рынке специалистов-посредников, профессионально занимающихся обслуживанием конечных пользователей, не только полезно, но и необходимо.

В качестве информационных посредников (брокеров) на информационном рынке выступают как специальные самостоятельные организации (это могут быть также информационные отделы организаций и библиотеки), так и частные лица, действующие на коммерческой основе.

## 4. Классификация баз данных

Под базой данных будем понимать упорядоченную совокупность информационных соединений (библиографических описаний опубликованных и неопубликованных источников, полных текстов и т. п.), относящихся к определенной теме и представленных на машинных (машиночитаемых) носителях — магнитных лентах, дисках, дискетах, компактных оптических (магнитооптических) дисках в виде специальным образом организованного файла или группы файлов [4, 12].

Первые машиночитаемые базы данных появились как побочный продукт, формируемый в ходе подготовки реферативных изданий, еще в 60-х годах. С запаздыванием примерно в 10 лет они появились в нашей стране и представляли собой аналоги печатных реферативных журналов, выпускаемых ВИНТИ (эти машиночитаемые базы готовились и распространялись на магнитных лентах, а записи были представлены в формате фотонаборной машины "Дигисет"). Это были базы, содержащие библиографические описания опубликованных источников — так называемые библиографические базы данных. К 1980 г. число разнообразных баз данных достигло 400 [4, 10], и в дальнейшем наблюдался устойчивый рост числа баз данных, производимых в мире. В настоящее время количество производимых в мире баз данных оценивается цифрой 8 тысяч, а годовой прирост числа баз данных превышает 10 %. С одной стороны, это дает возможность констатировать появление на рынке производителей новой мощной индустрии — индустрии баз данных, а с другой, обостряет проблему типологии баз данных. В то же время качественные характеристики, отличительные особенности и применимость базы данных (или группы баз данных) для решения конкретной задачи (задач) пользователя не могут быть оценены без знания типа базы данных.

Однако задача классификации уже существующих на информационном рынке (и возможных в будущем) баз данных не так проста, какой она представляется на первый взгляд. Это связано с тем, что глубоких исследований по проблемам построения некоей общепринятой типологии баз данных не существует, а сложившаяся зарубежная терминологическая практика (в приложении к типам баз данных) далеко не всегда адекватно отражает истинную типологию, и к тому же производители баз данных и вендоры не всегда следуют данной практике. Эта ситуация всесторонне исследована Н. И. Гендиной в [11], где описаны девять различных вариантов классификации баз данных. В табл. 1 приведена типология, принятая в различных системах классификации (в ней учтена синонимия в названиях типов баз данных и указывается только одно из синонимичных названий). Видно, что несмотря на достаточно большое количество типов, все их множество может быть условно поделено на два класса по характеру используемой для создания баз данных информации: на первичные, куда включаются числовые, текстовые, полнотекстовые базы данных и базы по физико-химическим свойствам, и вторичные, куда входят библиографические,

адресные, справочные и тому подобные базы данных.

Т а б л и ц а 1

Тип базы данных	Система классификации								
	1 [13]	2 [14]	3 [15]	4 [16]	5 [17]	6 [18]	7 [19]	8 [20]	9 [21]
Числовая	+				+	+			
Тексточисловая	+			+	+	+			
По свойствам	+								
Полнотекстовая	+		+	+	+	+			+
Библиографическая	+	+	+	+		+			+
Адресно-справочная	+								
Справочная		+	+р	+	+	+			
Цифровая		+		+					
База программ (программы для ЭВМ)		+		+	+				
Смешанная		+							
Адресная			+						
Фактографическая		+р					+	+	+р
Смешанная тексточисловая			+						
Библиографически-справочная				+р					
Условно-фактографическая				+					
Реферативно-библиографическая					+				
Реферативная						+			+
Тип БО (библиографическое описание)							+		
Тип БК (библиографическое описание и ключевые слова)							+		
Тип БКР (библиографическое описание и ключевые слова и рефераты)							+		
Тип БРШ (библиографическое описание и рубрикационные шифры)							+		
Тип БР (библиографическое описание и реферат/аннотация)							+		
Документальная							+р	+	
Документально-фактографическая								+	
Документографическая									+р
Лексикографическая									+

Примечание: р — тип базы данных в указанной системе классификации используется как родовое понятие.

С другой стороны, базы данных иногда подразделяют по типу хранимых в них сведений, и тогда один класс базы данных называют документальным, включая в него библиографические, полнотекстовые, справочные и другие базы, содержащие документальную, текстовую информацию; второй класс в этом случае образуют фактографические базы, к которым относятся числовые, тексточисловые и справочные базы, где данные носят, как правило, числовой характер. Однако на практике часто бывает невозможно однозначно отнести базу данных к какому-то одному типу, поэтому в большинстве классификаций присутствуют в той или иной форме смешанные базы данных.

Числовыми (или цифровыми) базами называют базы данных, содержащие структурированную информацию числового характера. В случае, если наряду с числовой информацией в базе данных содержится также и текстовая информация, базу называют тексточисловой. К полнотекстовым относят базы, содержащие полные тексты информационных источников (например, полные тексты патентов, законодательных актов, монографий и т. п.).



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
13	Информация для внешнеэкономической деятельности	8	7				1							
14	Спрос, реклама, коммерческие предложения отечественных и зарубежных фирм	21	16	2			1	2						
15	Ценовая информация	6	3		3									
16	Нормативно-правовая, юридическая информация	47	2	2				1	37	1			1	3
17	Деловые новости, дайджест прессы	5	2					2						1
18	Персоналии в политике, экономике, науке, технике	6	5											1
19	Статистика и экономика отраслей	10	3		2		3				1			1
20	Продукция, оборудование, технологии машиностроения	22	10				10	1					1	
21	Продукция и технологии приборостроения	22	5				15	1					1	
22	Оптика и оптические технологии	12	1				4						7	
23	Продукция, оборудование, технологии прочих отраслей	21	5			1	4	2					9	
24	Научно-техническая информация, патенты, изобретения	71	3				1	35	4	6			16	6
25	Вещества и материалы	28	5		1		13	1			7		1	
26	Вычислительная техника, электроника, программное обеспечение, телекоммуникации	32	14			11	1					1	3	2
27	Науки о земле, месторождения и запасы	22	6		1		4	4				1	3	3
28	Медицинская, биологическая информация	19	2			7				3		1	6	
29	Экология и охрана окружающей среды	13				1		4			1		2	5
30	Природоведение, сельское, лесное, рыбное хозяйство	21	2			2	10	1	1		1		3	1
31	Образование, обучение	11	5			3						1	2	
32	Библиотечное и архивное дело, литература	36	2				4	9	2	14			4	1
33	Культура, искусство, история, политика	29	7				7	2	3	4			4	2
34	Словари, классификаторы, тезаурусы, стандарты	34	1					1				32		
35	Сведения о базах данных и их владельцах	3	3											
36	Разное	8	1				6							1
	Итого	657	254	6	10	26	84	66	47	28	10	36	63	27

Примечание: Базы данных: 4 — справочные; 5 — полнотекстовые справочные; 6 — числовые справочные; 7 — документофактографические; 8 — фактографические; 9 — документографические; 10 — полнотекстовые; 11 — библиографические; 12 — числовые; 13 — лексикографические, словарные, словарно-справочные; 14 — тип БД не указан; 15 — смешанные.

## 5. Онлайн-базы данных и интерактивные службы

Машиночитаемая форма записи начала входить в информационную практику в начале 60-х гг. в применении к библиографической информации. За последующее десятилетие суммарный объем библиографических баз данных достиг 50 млн записей при ежегодном приросте от 4 до 8 млн библиографических описаний. Со второй половины 70-х гг. этот процесс охватил и другие виды информации — справочники, каталоги, отчеты и т.п., причем интенсивность этого процесса была такова, что в 1987 г. доля библиографических баз данных составляла лишь 41 % от всего числа баз данных, а остальные 59 % распределялись между так называемыми неблиблиографическими базами — числовыми, полнотекстовыми, справочными и т.п. [23, 24].

В период с 1982 по 1988 г. число баз данных выросло более чем в 5 раз, при этом количество записей (т.е. суммарный объем баз данных) увеличилось в 8 раз, а число ежегодно проводимых поисков составило в 1988 г. 26.5 млн против 7.5 млн в 1982 г., т.е. выросло более чем в 3.5 раза [4]. В 1995 г. в мире насчитывалось более 8 тыс. общедоступных баз данных, в том числе порядка 1 тыс. баз — полнотекстовых.

Параллельно с процессом роста числа, объемов и видов баз данных происходило формирование и развитие центров обработки баз данных и предоставления услуг диалогового доступа (*remote computing service vendors*) [5]. В начале 80-х гг. было несколько таких крупных центров (в литературе можно встретить также их другие названия — например, ИПС (информационно-поисковая система) [23], служба информационного поиска [4], Host-центр [25] и др.). Это Dialog, ORBIT, DATA-STAR, Mead Data Central, STN International, Telesystemes-Questel и др.

Крупнейшей службой информационного поиска является Dialog (США, г. Пало-Альто, Калифорния), принадлежавшая до 1988 г. аэрокосмической фирме Lockheed (США) и купленная затем американской компанией Knight Ridder [26]. Образованная в 1965 г., эта служба предлагала своим пользователям онлайн-доступ в 1979 г. к 86, в 1984 г. — к 250, в 1988 — к 274 базам данных [23, 27]. На сегодняшний день эта служба предоставляет в теледоступ более 400 разнообразных баз данных для более чем 155 тыс. абонентов в 100 странах мира [4], активно занимаясь также электронными изданиями — еще в 1987 г. она предоставляла пользователям 355 полнотекстовых изданий [23, 28].

Data-Star является крупной западноевропейской информационной службой, основанной в 1981 г. фирмой Radio Suisse Ltd (Швейцария) и предоставляющей пользователям 47 баз данных в 1984 г., более 100 баз данных — уже в 1987 г. [29], и 250 баз данных в 1995 г. [4]. В 1993 г. Data-Star стала европейским филиалом поглотившей ее американской службы Dialog.

Служба Mead Data Central (Дейтон, Огайо, США) была образована в 1968 г. Она предоставляет доступ к 450 базам данных и специализируется на полнотекстовой информации, новостях телеграфных агентств и т.п. В ее составе функционируют две системы, широко известные в мире, — LEXIS и NEXIS. Первая обеспечивает интерактивный доступ к полнотекстовой судебной-юридической информации США и ведущих европейских стран, а вторая — к полным текстам статей и других материалов из газет, относящихся к науке, технике, экономике, бизнесу, финансам, ценам, промышленному производству и т.д. В 1993 г. в системах LEXIS и NEXIS ежедневно осуществлялось до 200 тыс. поисков [4].

STN International (Scientific and Technical Information Network) объединяет информационные ресурсы трех служб — FIZ Karlsruhe (Центр информации по энергетике, физике, и математике, ФРГ), CAS (служба химической информации Chemical Abstracts Services,

США) и JICST (Центр научной и технической информации, Япония) и предоставляет доступ более чем к 170 базам данных.

Французская служба Telesystem Questel образована в 1979 г. и предоставляет пользователям более 150 баз данных, в том числе широко известную политематическую французскую базу данных Pascal. В 1994 г. она купила американскую службу ORBIT, где предоставлялось в теледоступ более 70 баз данных, и создала из этой службы свой североамериканский филиал [4].

Не останавливаясь далее на описании других известных служб (или центров обработки баз данных), заметим, что базы данных, предоставляемые в доступ различными службами, отчасти дублируются. Представление о степени дублирования баз данных (на примере ситуации в 1988 г.) дает табл. 3 [23].

Т а б л и ц а 3

Цифровое обозначение и название службы	Количество дублируемых БД											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) Dialog	—	59	34	51	25	16	5	4	22	10	10	10
(2) BRS	59	—	8	30	11	5	3	0	10	4	5	2
(3) ORBIT	34	8	—	15	13	16	4	0	2	0	1	1
(4) Data-Star	51	30	15	—	13	7	3	3	7	3	3	6
(5) ESA/IRS	25	11	13	13	—	6	4	3	1	0	1	1
(6) STN	16	5	16	7	6	—	1	0	0	0	0	1
(7) Questel	5	3	4	3	4	1	—	0	1	0	0	0
(8) Pergamon	4	0	0	3	3	0	0	—	1	0	0	1
(9) MDS	22	10	2	7	1	0	1	1	—	2	7	9
(10) Dow Jones	10	4	0	3	0	0	0	0	2	—	3	1
(11) VU/TEXT	10	5	1	3	1	0	0	0	7	3	—	4
(12) Textline	10	2	1	6	1	1	0	1	9	1	1	—

Дублирование баз данных выгодно как соответствующим службам ввиду их стремления увеличить прибыль, так и пользователям — к одной и той же базе данных пользователь может получить доступ в разных службах и волен выбрать наиболее выгодный для себя вариант. Например, 38 крупнейших баз данных в области финансов предоставляют в теледоступ 13 из 25 отраженных в справочнике [30] информационных служб (табл. 4).

В таблице приняты следующие обозначения.

Базы данных: 1 — ABC Germany; 2 — American Business Directory; 3 — American Business 20 Plus Companies; 4 — Cancorp Canadien Corporation; 5 — Company Intelligence; 6 — D&b Asia-Pasific Dun's Market Identifiers; 7 — D&B Canadien Dun's Market Identifiers; 8 — D&B Dun's Financial Records Plus; 9 — D&B Dun's Market Identifiers — US; 10 — D&B European Dun's Market Identifiers; 11 — D&B Eastern Europe; 12 — D&B Europe; 13 — D&B Israel; 14 — Directory of Directors; 15 — Disclosure Database; 16 — Disclosure / Worldscope; 17 — Extel Cards; 18 — FT Corporate Survey; 19 — FT Analysis Reports; 20 — Hoppenstedt Austria; 21 — Hoppenstedt Directory of German Company; 22 — ICC Company Reports & Accouts; 23 — ICC British Company Financial Datasheets; 24 — IMSWorld Pharmaceutical Company Profiles; 25 — Infocheck; 26 — Jordans Company Information; 27 — Kompass Asia / Pasific; 28 — Kompass Canada; 29 — Kompass Europe; 30 — Kompass UK; 31 — Media General Plus; 32 — Moody's Corporate News International; 33 — Moody's Corporate Profiles; 34 — S&P

Corporate Descriptions; 35 — S&P Daily News; 36 — S&P Register — Corporate; 37 — SEC Online; 38 — Teikoku Databank.

Информационные службы: A — Blaise — Line; B — CDP Online; C — Data — Star; D — Datastream; E — Dialog; F — Dow Joes News / Retrieval; G — Dusprint; H — ESA; I — Foodline; J — FT Profile; K — Global Scan; L — ICC Direct; M — IFR Securities Data; N — Infocheck; O — Infomart — Dialog; P — Jordans; Q — Kompass; R — MAID; S — NewsNet; T — Nexis; U — Questel-Orbit; V — Reuter Business Briefing; W — STN International; X — Waterlow Signature; Y — Wilsonline.

Т а б л и ц а 4

БД	Информационная служба																									
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	
1			*																							
2					*																*					
3					*																*					
4			*		*										*						*					
5			*		*																*					
6			*		*																*					
7			*		*																*					
8					*	*													*							
9			*		*	*													*							
10			*		*	*													*							
11			*		*	*													*							
12			*		*	*													*							
13			*		*	*													*							
14					*	*									*						*					
15			*		*	*													*		*					
16					*	*													*		*					
17			*		*	*				*									*		*					
18					*	*				*					*				*		*					
19			*		*	*				*									*		*					
20			*		*	*				*									*		*					
21			*		*	*				*									*		*					
22			*		*	*				*	*								*		*					
23			*		*	*				*	*								*		*					
24			*		*	*				*	*								*		*					
25			*		*	*				*	*			*					*		*					
26			*		*	*				*	*			*		*			*		*				*	
27					*	*				*	*							*		*						
28					*	*				*	*				*			*		*						
29					*	*				*	*						*		*		*					
30					*	*				*	*						*		*		*					
31					*	*	*			*	*						*		*		*					
32					*	*	*			*	*						*		*		*					
33					*	*	*			*	*						*		*		*					
34					*	*	*			*	*						*		*		*					
35					*	*	*			*	*						*		*		*					
36					*	*	*			*	*						*		*		*					
37					*	*	*	*		*	*						*		*		*					
38			*		*	*				*	*						*		*		*					

В табл. 5 приведено распределение баз данных, отраженных в североамериканском справочнике производителей онлайн-баз данных и информационных служб [31], по основным тематическим направлениям.

Т а б л и ц а 5

Тематика	Кол-во баз данных
1	2
Аэронавтика и космическое пространство	36
Сельское хозяйство и агротехнологии	77
Гидрология и океанология	17
Архитектура	18
Искусство	20
Астрономия	11
Авиация	33
Банковское дело	87
Библиография	17
Биология и биомедицина	97
Управление и бизнес	128
Химия и химические технологии	131
Коммерция	19
Продукты и товары	71
Коммуникации	82
Компьютеры и системы обработки данных	135
Конструирование	39
Потребительские товары	29
Корпорации	55
Курсы обмена валюты	23
Демография	61
Экономика	229
Образование	55
Электроника	38
Работа и карьера	39
Энергетика	125
Инженерия	54
Развлечения	23
Охрана окружающей среды	93
Финансы	256
Продукты питания и науки о питании	30
Лесоводство	26
Фонды и гранты	18
Геология	35
География	21
Правительственная информация	159
Здоровье и здравоохранение	67
Промышленность и промышленная продукция	103
Информационные системы и службы	98
Страхование	31
Языки и лингвистика	18
Законодательство	94
Библиотечные и информационные науки	46
Производство	56
Торговля и торговые рынки	145
Медицина	59

1	2
Металлургия	47
Музыка	5
Новости	124
Ядерные исследования	37
Патентная информация	46
Нефть	80
Фармацевтика	37
Физика	29
Математика	13
Политика и политология	22
Загрязнение окружающей среды	46
Психология	18
Издательское дело	42
Исследовательские проекты	41
Наука и технологии	113
Социальные науки	51
Социальные службы	22
Спорт	24
Стандарты и спецификации	22
Налогообложение	37
Токсикология	55
Туризм и путешествия	32

Сопоставим цифры, приведенные в табл. 5 и отражающие большую часть мирового информационного рынка в 1987 г., с аналогичными данными только одной службы — STN International по состоянию на 1995 г. [32] (табл. 6), и увидим, что эти цифры кардинально не отличаются.

Т а б л и ц а 6

Тематика	Количество баз данных
Сельское хозяйство	21
Биологические науки	41
Бизнес	22
Химия и химические технологии	39
Компании	26
Компьютеры и компьютерные системы	15
Инженерия	52
Охрана окружающей среды	43
Геология	17
Здоровье и здравоохранение	34
Материалы и материаловедение	35
Медицина	27
Патентная информация	18
Нефть	17
Фармакология	34
Физика	17
Токсикология	28

Со второй половины 80-х гг. заметно стала меняться внутривосточная ситуация в производстве баз данных: эта сфера перестала быть прерогативой крупных государственных организаций. Нарастающее же со временем число относительно небольших негосударственных образований, выступающих на информационном рынке в качестве генераторов

баз данных, в основном стало работать в области деловой и коммерческой информации и за короткий срок достигли в этой области заметных результатов, о чем свидетельствует анализ выпущенного в 1992 г. справочника [33] популярных баз данных деловой и коммерческой информации (табл. 7).

Т а б л и ц а 7

Тематический раздел	Количество БД	
	иностран.	отеч.
1	2	3
1. Сектор деловой информации		
1.1. Биржевая и финансовая информация		
1.1.1. Биржевая информация	7	3
1.1.1.1. Информация по фондовому рынку		
1.1.1.1.1. Информация по акциям.	4	—
1.1.1.1.2. Информация по облигациям	5	—
1.1.1.2. Информация по товарным рынкам	10	11
1.1.2. Финансовая информация		
1.1.2.1. Информация по обменным курсам валют	3	—
1.1.2.2. Информация по рынку денежных ресурсов	7	—
1.1.3. Комплексная биржевая, финансовая, экономическая и коммерческая информация	4	—
1.2. Экономическая и демографическая статистическая информация	31	18
1.2.1. Экономическая статистика		
1.2.2. Демографическая информация	3	7
1.2.2.1. Статистика спроса, потребителей и потребления	3	—
1.3. Коммерческая информация		
1.3.1. Информация по предприятиям и организациям		
1.3.1.1. Информация по промышленным предприятиям	16	53
1.3.1.2. Информация по предприятиям непромышленной сферы	7	13
1.3.1.3. Информация по научно-исследовательским организациям	3	4
1.3.1.4. Информация о финансовом состоянии и операциях предприятий и организаций	11	—
1.3.1.5. Информация по руководителям и специалистам предприятий и организаций	2	1
1.3.2. Информация по продукции		
1.3.2.1. Информация по промышленной продукции	19	14
1.3.2.1.1. Вычислительная и коммуникационная техника, программное обеспечение, электронные компоненты	16	17
1.3.2.1.2. Строительные объекты и технологии, недвижимость	3	4
1.3.2.2. Информация по продукции и услугам непромышленной сферы	6	6
1.3.2.2.1. Научно-технические результаты	4	17
1.3.3. Информация по коммерческим предложениям, инвестициям, контрактам	8	26
1.3.3.1. Электронные биржи	3	9
1.4. Деловые новости		
1.4.1. Общеэкономические деловые новости	14	18
1.4.2. Узкоотраслевые деловые новости	40	8
2. Сектор информации для специалистов		
2.1. Профессиональная информация		
2.1.1. Информация по медицине и фармакологии	5	—

1	2	3
2.1.2. Информация в области бизнеса и предпринимательской деятельности	—	21
2.1.3. Правовая информация	1	25
2.1.4. Информация в области электроники, вычислительной техники и информатики	6	10
2.1.5. Охрана окружающей среды	4	7
2.2. Научная и техническая информация для бизнесменов и предпринимателей	41	12
3. Сектор массовой и потребительской информации		
3.1. Электронные новости	9	22
3.2. Электронная литература		
3.2.1. Электронные энциклопедии	2	—
3.2.2. Электронные словари	3	13
3.2.3. Сведения о лицах и биографии	4	7
3.3. Потребительская и развлекательная информация		
3.3.1. Системы электронных покупок, банковских операций, обмена и рекламы		
3.3.2. Информация в области образования и трудоустройства	2	1
3.3.3. Потребительские справочники и новости	5	11
4. Сведения об адресах ведущих информационных фирм	83	185

За 12 лет с 1982 по 1994 г. рынок онлайн-баз данных существенно вырос. Так, число предоставляемых пользователям баз данных возросло более чем в 10 раз — с 773 баз в 1982 г. до 8261 в 1994 г. Количество вендоров (хостов) за этот же период увеличилось с 311 до 1629 (то есть в 5,2 раза), а число поисков, проведенных пользователями в онлайн-базах данных, выросло почти в 7 раз — с 7,5 до 51,78 млн [30].

## 6. Подбор баз данных, отвечающих запросу пользователя

Приведенные выше цифры наглядно свидетельствуют о том, что перед пользователем, желающим провести поиск информации по интересующей его проблеме, стоит ряд серьезных задач, и в том числе одна из первых — это выбор баз данных, охватывающих тематику его информационной потребности.

Для решения этой задачи в большинстве крупных систем реализован специальный режим поиска (назовем его расширенным или expand-поиском), который дает возможность пользователю обратиться к общему (для всех баз данных выбранной информационной службы или для указанного подмножества (кластера) баз данных) словарному файлу (часто это так называемый Basic Index, содержащий слова из полей заглавий работ, ключевых слов и др.), задав один или несколько связанных основополагающих для данного запроса терминов и получив в ответ перечень баз данных с указанием количества отвечающих заданным терминам документов (записей) в каждой базе.

Далее пользователь может осуществить выбор наиболее «представительных» баз данных и затем искать уже в них, используя все штатные средства поиска и вывода найденной информации. При этом пользователь может осуществлять поиск, не переключаясь с базы на базу последовательно, а сразу в выбранном тематическом подмножестве баз данных (кластере).

Но в реальной практике часто бывает так, что для обеспечения максимальной полноты информации, особенно при многоаспектном поиске, ресурсов одной службы недостаточно, и для получения пусть даже относительно небольшой доли (но, возможно, содержащей важные для пользователя сведения) дополнительной информации, необходимо привлечение ресурсов других информационных служб. Но тогда возникают сложности с освоением не одной, а нескольких информационно-поисковых систем, переформулированием запросов и т. п. В связи с этим на информационном рынке появились универсальные информационные и интеллектуальные шлюзы, представляющие собой специальные программные системы. Информационные шлюзы нацелены на облегчение доступа пользователя к ресурсам (базам данных) различных информационных служб и не предусматривают развитых средств помощи пользователю в выборе стратегии поиска и оценке результатов поиска. В интеллектуальных же шлюзах реализуются алгоритмы выбора баз данных, задачи формирования поисковых предписаний, а также ввод запросов на соответствующем языке, для чего используются методы искусственного интеллекта и элементы экспертных систем [33]. Различные шлюзовые службы могут охватывать ресурсы двух и более центров обработки данных [29], а службы Easy Net, Easylink, Searchline и InfoMaster являются службами с мировой сферой деятельности [29].

## 7. Перспективы Российского информационного рынка

Производство баз данных в России было начато примерно с середины 70-х гг. (имеются в виду промышленно тиражируемые по заказам организаций — потребителей баз данных). В настоящее время число созданных в стране баз данных составляет около 30 тысяч, в том числе доля крупных (более 100 тысяч записей) составляет 26 %, средних — 49 % и малых (менее 1 тысячи записей) — 25 %. При этом количество баз, содержащих массовую, коммерческую, официальную и финансовую информацию, не превышает 5 % от всего числа существующих баз данных. На мировом же информационном рынке большая часть центров-генераторов баз данных занята именно в области деловой и коммерческой информации [5]. Поэтому в ближайшие годы следует ожидать расширения и развития в России именно этого сектора информационного рынка.

По данным, приведенным в работе [5], в настоящее время сетями диалогового телекоммуникационного доступа охвачено немногим более 10 тысяч российских пользователей, причем в их числе около одной трети составляют иностранные пользователи, находящиеся в России. По-видимому, эта цифра никак не отражает сегодняшней потенциальной потребности пользователей в выходе на мировые и отечественные сети ЭВМ и базы данных, и уже в ближайшие годы следует ожидать резкого увеличения числа организаций и частных лиц, подключенных к различным сетям передачи данных, и прежде всего к сетям, дающим возможность выхода в мировую сеть Интернет.

В последнее десятилетие западный информационный рынок интенсивно наполняется базами данных на компактных оптических (CD-ROM) дисках. На сегодняшний день число производимых баз данных и мультимедиа дисков превышает 16 тысяч, причем наблюдается стабильный ежегодный рост количества производимых баз данных и фактически определились основные фирмы-производители. В России же производство CD-ROM баз данных пока находится в зачаточном состоянии, однако в соответствии с мировой тенденцией следует ожидать активного развития этой области деятельности.

Обилие информационных ресурсов, ставших потенциально доступными российскому пользователю, резко актуализирует проблемы их рационального и эффективного исполь-

зования, разумного сочетания возможностей приобретения баз данных на компактных оптических дисках, телекоммуникационного доступа к платным ресурсам host-центров и использования бесплатных ресурсов в сети Интернет. В связи с этим возрастает роль информационных брокеров и вновь на повестку дня ставится задача корпоративного использования дорогих информационных ресурсов и создания систем информационного обслуживания, работающих на принципах информационной кооперации.

## Список литературы

- [1] ПРАЙС Д. Малая наука, большая наука. В *“Наука о науке”*, Прогресс, М., 1966, 281–385.
- [2] МИХАЙЛОВ А. И., ЧЕРНЫЙ А. И., ГИЛЯРЕВСКИЙ Р. С. *Научные коммуникации и информатика*. Наука, М., 1976.
- [3] СОЛТОН ДЖ. *Динамические библиотечно-информационные системы*. Мир, М., 1979.
- [4] АРСКИЙ Ю. М., ГИЛЯРЕВСКИЙ Р. С., ТУРОВ И. С., ЧЕРНЫЙ А. И. *Инфосфера: информационные структуры, системы и процессы в науке и обществе*. ВИНТИ, М., 1996.
- [5] АРСКИЙ Ю. М., ГИЛЯРЕВСКИЙ Р. С., ЕГОРОВ В. С. и др. *Информационный рынок в России*. ВИНТИ, М., 1996.
- [6] РОДИОНОВ И. И. *Основные характеристики мирового рынка информационных услуг*. МЦНТИ, М., 1990.
- [7] КАДОМЦЕВ Б. Б. Об обеспечении научно-технической информацией учреждений Академии наук СССР. *Вестн. АН СССР*, 1981, 12–16.
- [8] КУЛИК А. Н. *Информационные сети и языковая совместимость дескрипторных информационно-поисковых систем*. Сов. радио, М., 1977.
- [9] The knowledge. *Economist*. November 11th, 1995, 75.
- [10] КУАДРА К. Исторический обзор развития и использования баз данных. В *“I конф. “Восток — Запад” по научной, технической и деловой онлайн-информации”*, Доклады, М., 1990, 15–28.
- [11] ГЕНДИНА Н. И. *Лингвистическое обеспечение библиотечной технологии: Дис. ... докт. пед. наук*. СПб., 1994.
- [12] ЯКУБАЙТИС Э. А. *Информационные сети и системы: Справочная книга*. Финансы и статистика, М., 1996.
- [13] ЧЕРНЫЙ А. И. Роль информационных ресурсов в развитии национальной экономики. *Вопросы информационной теории и практики*, №57, 1989, 15–62.
- [14] СЕЙФУЛЬ-МУЛЮКОВ Р. Б., АЛЕКСАНДРОВА Н. В. Информационный рынок капиталистических стран (Обзор). *НТИ*, сер. 1, 1989, 24–27.
- [15] МИХНЕВИЧ К. П. Автоматизированные информационные системы США и Западной Европы: Обзор зарубежных источников. *Там же*, №2, 1989, 23–29.

- [16] РЕБРОВА М. П., ЭНГЕЛЬГАРТ А. В. О "Справочнике баз данных, доступных в режиме диалога" (США). *Там же*, №5, 1988, 22–27.
- [17] *Базы данных стран — не членов МСНТИ по тематике приоритетных направлений КП НТП: Справочник*. МЦНТИ, М., 1988.
- [18] *Базы данных стран Латинской Америки и Карибского бассейна: Справочник*. МЦНТИ, М., 1990.
- [19] *Каталог баз данных на магнитных лентах, распространяемых органами научно-технической информации в ГАСНТИ*. ВИНТИ, М., 1986.
- [20] *Каталог локальных баз данных научно-исследовательских учреждений и организаций АН СССР и АН союзных республик*. ВИНТИ, М., 1990.
- [21] *Базы данных в СССР: Каталог 2100 описаний баз данных по всем отраслям науки, техники и народного хозяйства*. ВИНТИ, М., 1990.
- [22] *Российская энциклопедия по информации и телекоммуникации. Ч. 3. Базы данных. Справочные издания. Программные продукты. Телекоммуникационные системы и их информационные ресурсы*. МБИТ, М., 1994.
- [23] ЩЕРБАТОВА А. К. Индустрия баз данных и интерактивных информационных служб. *Итоги науки и техники, сер. Информатика*, **14**, 1990, 179–209.
- [24] ВЕРЕВЧЕНКО А. П. Создание машиночитаемых информационных ресурсов — одно из условий интенсификации научно-технического прогресса. *Прикладная информатика*, №4, 1988, 5–15.
- [25] *Телекоммуникационные сети и их информационные ресурсы*. МБИТ, М., 1994.
- [26] МИХНЕВИЧ К. Продается система Dialog. *ЭИ, сер. Информатика*, №9, 1989, 10–11.
- [27] Multidisciplinary services. *Libr. Technol. Reports*, **24**, No 3, 1988, 313–319.
- [28] PAGELL R. A. Searching full-text periodicals: how full is full? *Database*, **10**, No 5, 1987, 33–36.
- [29] РОДИОНОВ И. И. *Основные характеристики рынка информационных услуг развитых капиталистических стран*. МЦНТИ, М., 1990.
- [30] *The Online Manual. A Practical Guide to Business Databases*. S. Allcock and J. Osborne (eds.), Learned Information Ltd, 1994.
- [31] *The North American Online Directory*. Bowker Company, N. Y., 1987.
- [32] *STN International. Databases in Science & Technology*. FIZ, Karlsruhe, 1995.
- [33] РОДИОНОВ И. И. *Мировой рынок электронной информации для предпринимателей. Ч. 2. Справочник популярных баз данных деловой и коммерческой информации. Сер. Технологии электронных коммуникаций*, **22**, М., 1992.