

## **АНАЛИЗ СПРОСА ДЕПОЗИТАРНЫХ ФОНДОВ ГПНТБ РОССИИ И РОЛЬ ЭЛЕКТРОННЫХ ТЕХНОЛОГИЙ**

Г.А. Евстигнеева, А.И. Земсков, ГПНТБ России, 101999, г. Москва, ул. Кузнецкий мост, 12, fo2@gpntb.ru, zemskov@gpntb.ru

### **DIGITAL FORMAT FOR DEPOSITORY PART OF SCIENTIFIC PUBLICATIONS' COLLECTION**

Andrei Zemskov, Galina Evstigneeva, Russian National Public Library for Science and Technology, Moscow, Russia, fo2@gpntb.ru, zemskov@gpntb.ru

As a rule, any large library has so called depository part of collection. Specific circulation of these publications is not high, but preservation, conservation and provision of access to given collection need permanent efforts. The paper presents evaluation of expenses and centralized systems efficiency. These systems (e.g. JSTOR, Backfiles for Future) are publicly accessible collections of full text versions of scientific periodicals.

#### **Роль и место электронных публикаций в фондах научных библиотек**

Хорошо известно, что использование электронных публикаций дает возможность получения информации на 2-3 месяца раньше, чем в печатном виде, позволяет обслуживать большее число пользователей, предоставляет более удобный поиск информации, а в ряде случаев дает экономию средств (использование бесплатных информационных ресурсов, оформление доступа на условиях консорциума). В докладе мы хотим обратить внимание на использование электронных коллекций, способных взять на себя функции депозитариев.

Некоторые библиотеки для создания электронных архивов пошли по пути оцифровки изданий фонда собственными силами. Это очень серьезная работа. Она требует разработки концепции создания электронной библиотеки, включающей подходы к отбору литературы для сканирования, выработку технологии, стандартов, обеспечение доступа к информации. Организация и проведение работ связано со значительными финансовыми расходами и немалыми дополнительными трудозатратами. При этом вопрос авторского права тоже нельзя оставить без внимания. Поэтому сегодня мы будем говорить лишь о коллекциях, предлагаемых рынком электронных информационных продуктов.

Практически все ведущие зарубежные научные издательства и научные общества представляют свои издания в электронном виде. В первую очередь это касается научных журналов за текущие годы. Так как эта работа началась несколько лет тому назад, то и электронные коллекции в основном имеют ретроспективу за 5-7 лет. Как правило, при покупке лицен-

зии библиотека получает доступ к электронной версии журнала за все эти годы. Для ученого или специалиста, занимающегося изучением современного состояния науки, такой охват вполне достаточен. Однако потребность в литературе за более старые годы тоже существует. Поэтому библиотека обязана заботиться о полноте коллекции. В определенной степени здесь могут использоваться электронные ресурсы. Речь идет о наиболее значимых изданиях., которые входят в состав электронных архивов, создаваемых издательствами, научными обществами, различными некоммерческими объединениями пользователей и распространителей и формируемых по тематическому принципу. Вот некоторые примеры.

Проект JSTOR (Journal Storage, электронный архив научной периодики) [1], некоммерческая организация, был инициирован фондом Эндрю Меллона, сейчас поддерживается также фондом Буша и реализуется университетом штата Миннесота и рядом частных колледжей. Первоначальная задача – за счет конвертирования печатных малоспрашиваемых периодических изданий в электронный формат создать условия для долговременного и компактного хранения научной периодики, добиться экономии библиотечных площадей и расходов на содержание библиотечных депозитарных фондов. Удобные поисковые средства превратили систему из просто архивной в мощный инструмент научной работы и обучения. Экономическая суть идеи состоит в том, что библиотеки, входящие в JSTOR на корпоративных условиях, занимаются оцифровкой для предоставления пользователям хорошо продуманной тематической коллекции за умеренную гибкую плату, что дает экономическую выгоду как каждой библиотеке-участнице проекта, так и пользователям баз данных.. Еще одна особенность проекта, которую могут взять на вооружение библиотеки, начинающие создавать собственные коллекции, – введение понятия и практики применения так называемой «подвижной стенки» (moving wall), то есть фиксированного промежутка времени (в большинстве случаев от 2 до 5 лет) между временем публикации данного тома и его появлением в коллекции JSTOR. Такой подход позволяет издателям сохранить в основном свой сектор рынка, свои доходы и привлекает их к сотрудничеству с JSTOR. О разумности политики JSTOR свидетельствует то, что с библиотекой JSTOR сотрудничают 156 издательств, в основном университетских, выпускающих 290 наименований журналов. В фонде имеется 700 000 статей (более 9 млн. страниц), к фонду обращаются 1300 коллективных пользователей. С 1997 года, с момента основания JSTOR, по настоящее время распечатано 6,4 млн. статей. Библиотека JSTOR организована по принципу тематических коллекций по бизнесу, по экологии и ботанике, по фундаментальным наукам и т.д.). Определенной гарантией разумного решения всех технологических вопросов хранения и обслуживания в данном проекте является тот факт, что проект организован профессиональными специалистами, а не коммерсантами.

Издательства также понимают необходимость создания полных коллекций изданий. Интересным для библиотек может быть пример издательства Эльзевир, приступившего к созданию архивных коллекций журналов [2]. Проект разработан в рамках программы Science Direct . Он носит название «Архивы для будущего» (Backfiles to the Future). Этот проект предполагает оцифровку полных текстов ретроспективных выпусков в формате pdf всех 1200 наименований научных журналов, издаваемых и изданных ранее Эльзевир. Ретроспекция простирается вплоть до первого выпуска первого тома, то есть – в некоторых случаях , до ста лет давности. Всего в проекте будет оцифровано свыше 40 миллионов страниц. Первая тематическая группа (органическая химия, неорганическая химия, химическое машиностроение, физическая и аналитическая химия) уже готова. Во вторую группу включены журналы по математике, экономике, эконометрике и финансам, менеджменту и бухгалтерскому делу, психологии, общественным наукам, наукам о Земле, физике и астрономии. Завершать проект в 2003 году будут оцифрованные журналы по компьютерным наукам, медицине, сельскому хозяйству и биологии, экологии, материаловедению и энергетике. Доступ предполагается предоставлять при наличии одноразовой лицензии, дающей неограниченные во времени права поиска, просмотра и загрузки электронных материалов. С учетом меньшего спроса на архивные копии научных журналов и трудности обеспечения их физической сохранности, возможность получения материалов из такого рода источника весьма перспективна.

Из практики ГПНТБ России. Примерно половина журналов Эльзевир (свыше 600 наименований) соответствуют тематическому плану комплектования библиотеки. И практически все они поступали в фонд до 1991 года. Начиная с 1992 года поступление литературы (и журналов Эльзевир в том числе) сократилось в десятки раз. С 1997 года в России стала действовать Электронная библиотека (E-library) [3]), организованная и поддерживаемая Российским фондом фундаментальных исследований. Важное место в этой библиотеке занимают журналы Эльзевир. Но и она покрывает потребности ГПНТБ лишь на 35%. Кроме того, в E-library журналы Эльзевир представлены, начиная с 1994-1997 годов. Поэтому тематические архивные коллекции издательства представляют определенный интерес.

Нельзя обойти вниманием и проекты научных обществ по созданию электронных архивов. Научные общества традиционно занимались и занимаются издательской деятельностью, целью которой является создание наиболее благоприятных условий для работы ученых и исследователей. Электронные публикации имеют ряд неоспоримых преимуществ по сравнению с печатными. Поэтому научными обществами собраны ценные коллекции электронных публикаций и встал вопрос о создании электронных архивов, доступных ученым всего мира. В качестве примера можно привести инициативу Европейского математического общества. В 2001 году

был начат проект Electronic Mathematics Archives Network Initiative (EMANI) [4], направленный на создание математического электронного архива. В состав инициативной группы вошли:

The Cornell University Library, Ithaca, N.Y

The State and University Library Goettingen, (SUB Goettingen

The Tsinghua University Library, Beijing.

The Orsay Mathematical Library, Paris.

В проекте также участвуют ведущие научные издательства: Springer-Verlag, Birkhaeuser Verlag, Teubner Verlag, Vieweg Verlag, а также электронная библиотека ElibM, доступная через Европейскую Математическую Информационную службу (European Mathematical Information Service – EMIS). Эти четыре издателя имеют давние традиции публикаций по математике. В свою очередь ElibM является плодом сотрудничества нескольких журналов и редакторов на добровольной основе, предоставляющих электронные публикации для использования через Интернет.

В ближайшие годы проект будет развиваться, и количество партнеров и пользователей возрастет.

### **Оценки затрат на хранение депозитарных фондов**

Депозитарные фонды определяются достаточно условно, с учетом специфики данной библиотеки и состава фондов, а также возможностей их размещения и спроса на конкретную тематику. Обычно они составляют 50–75% от общих фондов крупной библиотеки, в ГПНТБ России депозитарные фонды составляют около 5 млн. экземпляров. Нужда в дополнительных помещениях – постоянная забота любой библиотеки, она рано или поздно возникнет, если не заниматься списанием и передачей литературы «на сторону».

Посмотрим, во что обходится налогоплательщику технология традиционного хранения. Эти данные приведены в табл. 2.

**Таблица 2.** Усредненные фактические данные ГПНТБ России по используемым для удаленного хранения помещениям

Практическая плотность хранения	672 экземпляра на кв. метр
Стоимость нового строительства	100 долларов за кв. метр
Эксплуатация имеющихся хранилищ	230 рублей (около 7,5 долларов) за кв. метр в год

Можно считать, что удельная стоимость хранения литературы за один год обходится библиотеке и налогоплательщику ориентировочно в 1000–1200 долларов за 100 тысяч единиц хранения или около 10 тысяч долларов за фонд в 1 миллион единиц хранения. Следовательно, годовая

стоимость хранения депозитарного фонда ГПНТБ России составляет примерно 50 тысяч долларов.

Заметим, что испытанные способы компактирования не решают проблему в принципе, а лишь отодвигают время ее реализации. Действительно, подвижные стеллажи (так называемое компактное хранение) снижают потребность в площади лишь в 2–3 раза, не более. Хранение микроформ, по нашим данным, дает экономию места в 15–30 раз по сравнению с обычным способом хранения. Если планировать перевод фонда на CD-ROM, то выигрыш будет еще заметнее, хранение на оптических дисках даже в формате изображения в несколько сотен раз компактнее, чем в обычном варианте. И все же неуклонное наращивание депозитарной части фонда – неизбежное свойство традиционной библиотеки. Принципиальное решение вопроса – либо вообще отказ от хранения (например, библиотека Института научно-технической информации Республики Корея комплекзует 20 тысяч наименований научно-технической периодики, а хранит издания только за 3 последних года), либо передача функций хранения в специализированную внешнюю организацию.

### **Мотивация хранения**

Следующая группа вопросов касается мотивации хранения. Если речь идет о национальных библиотеках или приравненных к ним (в том числе получателям обязательного бесплатного экземпляра), то все вопросы, пожалуй, отпадают. Для других библиотек (в первую очередь, научно-технических) достойной мотивацией (помимо общечеловеческих рассуждений об исторической роли библиотек в сохранении национального наследия) может быть заметный спрос как со стороны своих читателей, так и спрос на литературу от других библиотек по линии межбиблиотечного абонемена. В каком формате хранить – привычном печатном или в электронном – этот вопрос сейчас широко дискутируется. Дополнительный стимул дискуссии придает тот факт, что сейчас какая-то часть библиотечных поступлений существует только в электронном виде. Кроме того, от того, какое значение мы придаем депозитарному фонду и какова его востребованность, зависят решения, связанные с ретроспективной конверсией карточных каталогов. Известно, что существуют различные возможности для ретроконверсии каталогов, которые дают различные поисковые возможности в электронном каталоге. Наиболее простым и дешевым является сканирование карточки для воссоздания ее графического образа. Это наиболее дешевый способ, но и поисковые возможности при этом будут ограничены. Хранить информацию в виде графических образов карточек можно только в том случае, если спрос на издания депозитарного фонда крайне невелик. Сканирование с распознаванием и переводом карточки в текстовый формат (процессы, требующие обязательного редактирования), или ручной ввод в электронный каталог являются более трудоемкими и доро-

гостоящими способами. Они эффективны при частом обращении к электронному каталогу, то есть, если издания пользуются высоким спросом. Таким образом, решения библиотеки о депозитарном фонде влияют на представления о том, каким должен быть электронный каталог.

Оценки спроса на издание в зависимости от года публикации можно попытаться сделать и на основе теоретических рассуждений. Ясно, что всегда между моментом публикации и возможным максимальным спросом существует некий промежуток времени, в течение которого проходят технологические операции – движение издания от издательства в библиотеку, внутрибиблиотечная обработка издания (инвентаризация, каталогизация, индексирование) и расстановка его на полку. Протяженность этого периода составляет несколько месяцев. Затем начинается также латентный до поры до времени период интеллектуального ознакомления читателей с изданием и формирование спроса на него; реклама в издательствах, работающих с научно-технической литературой не очень агрессивна. Так что должно пройти минимум 12–20 месяцев от момента публикации и до проявления активного спроса на материал.

Еще одно теоретическое положение – корреляция года выпуска научного издания с научным стажем пользователя. За исключением классических изданий, сохраняющих познавательную ценность сквозь века и десятилетия, научные интересы студента начинают проявляться где то между 2-м и 4-м курсом ( то есть в возрасте 19–20 лет) и продолжают эти интересы всю активную научную жизнь. Чтобы не отягощать вас дополнительными рассуждениями о творческом долголетии отдельных корифеев науки (не они определяют среднюю величину спроса в библиотеках), условимся, что активный интерес к изучению периодики заканчивается с наступлением пенсионного возраста, то есть к 55–60 годам. Итак, в течение примерно 40–45 лет спрос отдельно взятого человека может в основном сохраняться, поскольку справочные и познавательные интересы индивидуума меняются не быстро. В какой то мере перекликаются с нашими теоретическими рассуждениями (и подтверждают их) данные Российского фонда фундаментальных исследований о распределении по возрастам грантополучателей РФФИ. По всем научным дисциплинам зависимость числа полученных грантов от возраста отражается ясно выраженными колоколообразными кривыми распределения, с резким ростом при достижении возраста 23, 42 и 50, 60 лет, и плавным спадом после 60–62 лет . Исходя из данных РФФИ, можно полагать, что 40-летний период между 20 и 60 годами может иметь отражение и в анализе спроса отдельно взятого ученого на научную литературу, то есть происходит нечто вроде резонансного отклика.

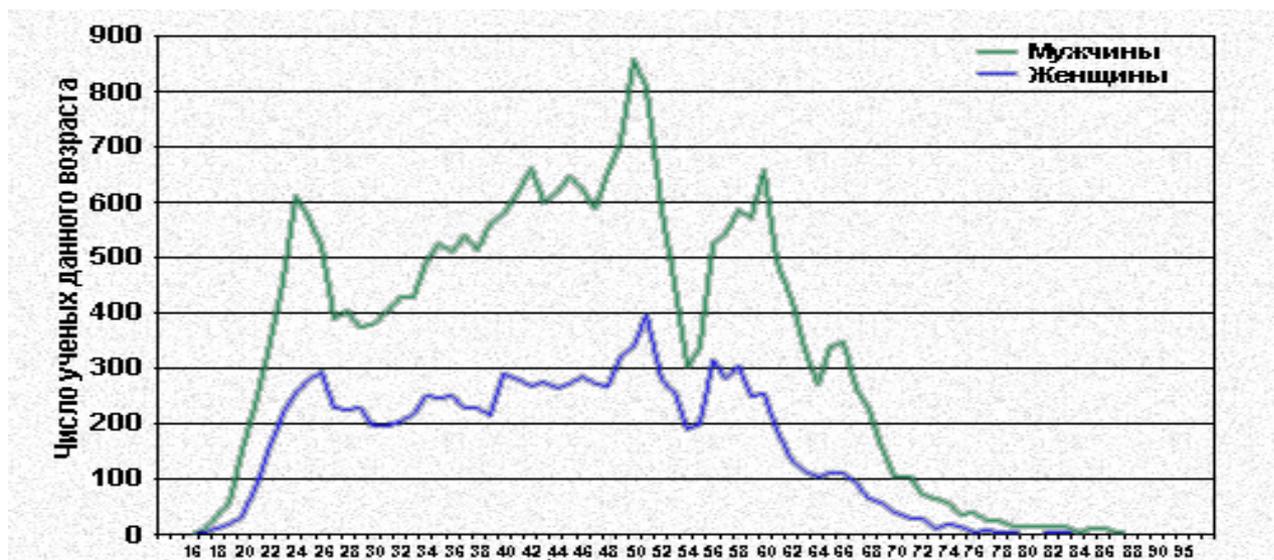


Рис. 1. Ученые — грантополучатели грантов РФФИ: распределение по возрасту

Достаточно интересными могут оказаться данные по запросам системы межбиблиотечного абонемента. В дополнение к выше рассмотренным факторам здесь существенную роль оказывают корреляция спроса с наличием какого-либо издания в низовой библиотеке: на предприятии, в научном учреждении или институте.

Фактические показатели динамики спроса на литературу.

Рассмотрим фактические данные по ГПНТБ России за 1997–2001 гг., представленные сотрудниками отдела хранения, отражающие долю спроса читателей читальных залов на депозитарную литературу (издания, опубликованные до 1990 года) в общем спросе (табл. 3).

Таблица 3. Доля спроса на депозитарную литературу, опубликованную до 1990 года; за 100% берется общее количество запросов. Наблюдения за 1997–2001 годы

Виды изданий	Процентное отношение к общему количеству запросов				
	1997	1998	1999	2000	2001
1. Книги, всего	19,9%	17,3%	18,4%	18,2%	18,0%
В том числе:					
Книги отечественные	15,5%	15,8%	17,0%	16,6%	16,8%
Книги иностранные	4,4%	1,5%	1,4%	1,6%	1,2%
2. Журналы, всего	9,3%	11,5%	8,1%	16,2%	16,3%
В том числе :					
Журналы отечественные	4,8%	6,9%	4,6%	5,2%	5,0%
Журналы иностранные	4,5%	4,6%	3,5%	11,0%	11,3%
3. Переводы	3,8%	4,3%	4,5%	2,3%	2,0%
4. Авторефераты	0,3%	0,6%	0,7%	2,4%	1,3%
5. Итого (сумма 1 + 2+ 3 + 4)	33%	33,0%	31,7%	39,1%	37,6%

Заметим, что колебания спроса от года к году носят достаточно случайный характер, поскольку реальная выдача за определенный год определяется многими факторами: текущими запросами и интересами пользователей, тиражами в год выпуска литературы, объемом библиотечного комплектования (например, ГПНТБ России пришлось неожиданно взять в свои фонды довольно крупную коллекцию неопубликованных переводов). Оказывает свое влияние и доступность или недоступность данной части фонда по техническим причинам, например, из-за закрытия данного хранилища ввиду наводнения или другой аварии, закрытием библиотек на ремонт или ввиду каких-то чрезвычайных обстоятельств (например, как наказание за невозможность вовремя оплатить счета за электроэнергию) – все это в ГПНТБ России происходило.

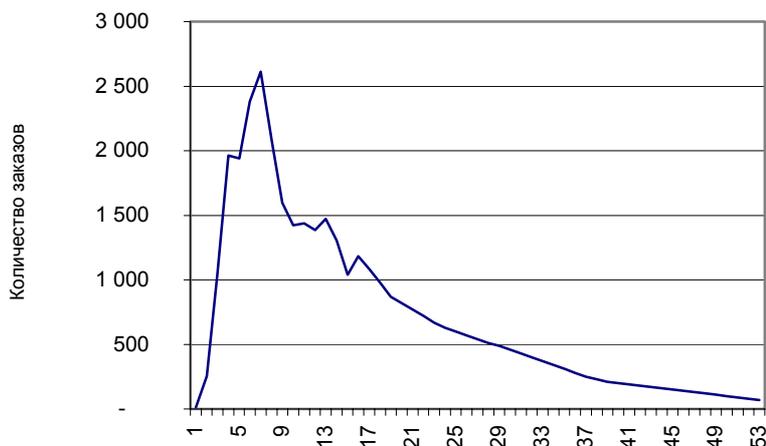
Оценки фактической спрашиваемости литературы удаленными пользователями в зависимости от года ее публикации мы получили также на основе данных автоматизированной системы Межбиблиотечного абонемента ГПНТБ России. Ввиду того, что в системе МБА много лет существует автоматизированный учет транзакций, нам удалось (спасибо сотрудникам ГПНТБ России Е.Б. Подмориной и Г.С. Лисицыной) получить более развернутую информацию. Динамика спроса в текущем, 2002 году, на литературу (заказы системы МБА), выпущенную в свет в указанном десятилетнем периоде, представлена на рисунке 2.



Полученные результаты представляются достаточно значимыми, поэтому мы провели анализ спроса в системе отечественного МБА и международного МБА за более длительный период наблюдения (изучен спрос за 5 лет, с 1997 по 2002 гг. включительно, на литературу, изданную, начиная с 1930 года).

Результаты показаны на рис. 3 (обращения отечественных библиотек) и рис. 4 (обращения зарубежных библиотек).

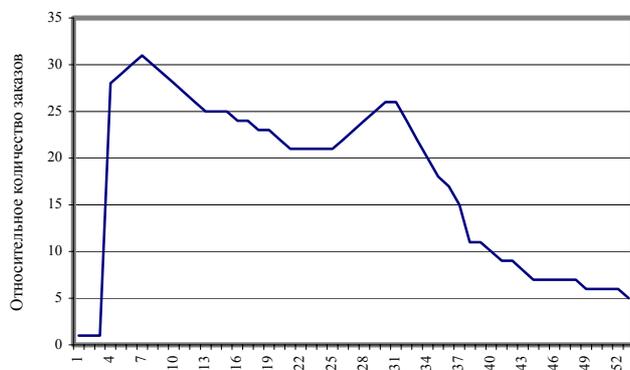
**Рис 3.** Количество заказов по системе межбиблиотечного абонеента в зависимости от года публикации издания



"Возраст" издания - количество лет, прошедших с момента публикации до момента наблюдения (1997 - апрель 2002 года)

Данные наблюдений спроса пользователей отечественных библиотек за последний год (рис. 2) и за много лет (рис. 3) не сильно отличаются; более тщательная проработка деталей на рис. 3 позволяет выделить следующие типовые черты : «инкубационный» период порядка 2-3 лет, затем довольно крутой подъем и максимум, относящийся к литературе 7-летней давности и затем плавное снижение спроса. Трактовка более тонких деталей спроса, относящихся к периоду начала 1990-х гг. и к 1998 году, уведет нас от научного в политическое рассмотрение, поэтому мы не комментируем известные факты. Запросы зарубежных библиотек по линии международного абонеента демонстрируют колоколообразную зависимость от времени с условной шириной «колокола» порядка 40 лет (рис. 4).

**Рис 4.** Относительное количество заказов по системе международного МБА в зависимости от года публикации издания



"Возраст" публикации - количество лет, прошедших с момента публикации до момента наблюдения (апрель 2002 года)

Можно обратиться к хорошо известному в библиотечных кругах правилу Парето, который гласит, что 80% спроса обеспечивается 20% публикаций, находящихся в фонде библиотеки. Применение этого подхода к нашим фактическим данным показывает, что и для обслуживания в читальных залах и для обслуживания по системам МБА и ММБА спроса обеспечивается литературой, изданной в промежутке, занимающем примерно 25 – 30 лет.

Итак, хранение и предоставление в обслуживание депозитарных фондов необходимо, но обходится недешево. И все же наиболее уязвимыми местами традиционного способа хранения в удаленных и не приспособленных хранилищах являются два фактора: гибель литературы по причине частых аварий теплосетей и сетей канализации и недопустимо большие сроки выполнения заказа на литературу, хранящуюся в этих подвалах. Гибель периодического издания отвратительна не только сама по себе, но и потому, что нарушается целостность всей коллекции данного наименования журнала. Что касается срока в 2 дня, которые требуются для доставки заказа в удаленное хранилище и выполнения его, то, конечно, это неприемлемые сроки для предоставления научно-технической информации в наше время и для библиотеки, имеющей выход в Интернет. Еще более обидными являются многонедельные паузы в работе хранилищ, вызванные какими-то авариями или катастрофами; на это время обслуживание литературой из данного хранилища не ведется.

#### Заключение

Не случайно в данной публикации представлен материал об электронных архивах и приведены данные по депозитарным фондам. Рассматривая эти два вопроса в комплексе, мы пытаемся решить проблемы сохранности фондов.

Видно, что спрос на литературу за последние 30 – 40 лет остается заметным и пренебрегать им никак нельзя.

При этом спрос в читальных залах индивидуальных пользователей на литературу 10-летней (и более) давности составляет, как мы уже отмечали, около 35%, а в структуре спроса по системе МБА литература 10 и более лет давности составляет значительно существенную долю, около 60 %. Это наблюдение легко объяснимо из самой сути МБА – обращаются за помощью в том случае, когда своих сил не хватает, а в новых научно-технических библиотеках откуда же взяться старым публикациям? Из Таблицы 3 мы видим, что спрос на литературу десятилетней и более давности составляет около 35% общего спроса, причем половина этого спроса падает на книги, в особенности, отечественные книги. Это наблюдение легко поддается теоретическому обоснованию: книги действительно являются существенно инерционным источником информации: они неспешно “вызревают”, поскольку именно в книгах аккумулируется устоявшееся, обоб-

ценное знание, но зато по той же причине они проявляют наибольшее “творческое долголетие”.

Многолетний опыт показывает, что попытка сохранить многомиллионный фонд в непригодных или даже аварийных помещениях, разбросанных по всему городу, имеет плачевные результаты. Литература гибнет или годами не выдается. Сроки доставки литературы читателю не выдерживают никакой критики. Поэтому нельзя пренебрегать возможностью использовать электронные научные тематические коллекции. Проекты, направленные на создание централизованных и специализированных систем хранения и предоставления в обслуживание полнотекстовых электронных копий малоспрашиваемых периодических изданий (проект JSTOR), или проект оцифровки всех публикаций данного издательства с момента выпуска первого номера и также предоставление их в открытый (платный) доступ (проект Backfile for the Future издательства Elsevier) нас, конечно, очень интересуют. Такой подход обеспечит нормальные, профессионально продуманные условия обеспечения сохранности изданий.

Конечно, мы не призываем к тому, чтобы «взять все книги, да и сжечь». Прежде чем перевести какую либо часть фонда на электронные носители, должна быть проведена серьезная всесторонняя комплексная оценка фонда, с одной стороны, и изучение рынка информационной электронной продукции, с другой. Кроме того, было бы неосторожно целиком и полностью довериться коммерческому посреднику, предоставляющему не информационный продукт, а лишь доступ к нему. Здесь имеется опасность потери информации в том случае, если она, как предмет торговли, станет недостаточно прибыльной. Поэтому, важно подписать лицензионное соглашение, позволяющее библиотеке иметь наиболее благоприятные условия. Кроме того, необходимо как можно шире использовать сотрудничество библиотек в консорциумах.

Данное сообщение является приглашением к обсуждению интересной на наш взгляд темы. Возможно, эта проблема заинтересует РФФИ и ее можно будет решить в рамках проекта E-library.

## Литература

[1] Домашняя страница JSTOR: <http://www.jstor.com>

[2] Домашняя страница Elsevier: <http://www.elsevier.com>

[3] Домашняя страница E-Library: <http://www.elibrary.ru>

[4] Бернд Вегнер: Электронная математическая библиотека Европейской службы математической информации – опыт распределенных недорогих издательств, 8-я Международная конференция «Крым 2001» Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества, том 1, стр. 317-320