

## **СОВРЕМЕННОЕ СОСТОЯНИЕ И ПЕРСПЕКТИВЫ РАЗВИТИЯ АВТОМАТИЗИРОВАННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ**

**В.А.ГЛУХОВ, О.Л.ЛАВРИК  
РАСПРЕДЕЛЕННЫЕ ЭЛЕКТРОННЫЕ БИБЛИОТЕКИ**

Библиотеки были изобретены человечеством, чтобы хранить документальные источники и предоставлять их в общественное пользование. Библиотеки и занимались комплектованием документов в формах, присущих каждой конкретной эпохе – глиняных табличках, папирусных свитках, рукописях, печатных изданиях и, наконец, изданиях на электронных носителях.

### **Библиотечные Интернет-технологии**

Развитие сети Интернет поставило перед каждой библиотекой задачи, которые условно можно разделить на две группы: 1) организация удаленного доступа к письменным источникам, хранящимся в библиотечных фондах; 2) обеспечение доступа для своих читателей к “чужим” информационным ресурсам, доступным через Интернет.

Если проанализировать развивающиеся сейчас направления библиотечных работ, связанных с Интернетом то можно заметить, что первая группа задач решается главным образом за счет:

- обеспечения доступа к собственным электронным каталогам, проблемно-ориентированным библиографическим базам данных, сводным каталогам библиотек, а также организации доступа к каталогам различных библиотек по сетевым поисковым протоколам типа Z39. 50;
- публикации в Интернете электронных версий специализированных библиотечных изданий (журналов, библиографических списков, указателей литературы, каталогов выставок и т.д.);
- создания полнотекстовых баз данных (включающих как издания самих библиотек, так и электронные версии книг и журналов из их библиотечных фондов) и обеспечения поиска в этих массивах через Интернет;
- предоставления услуг по электронной доставке документов и электронному межбиблиотечному абонементу.

Российские библиотеки и информационные центры получили возможность работы с Интернетом не так давно, примерно в 1996-1997 гг. Таким образом, теоретические и практические разработки в этой области, по сути, только начинаются. Прежде всего, они сосредоточены на обеспечении доступа к электронным каталогам и базам данных для удаленных пользователей библиотек. Это каталоги таких организаций, как БЕН (<http://ben.irex.ru>), ГПНТБ СО РАН (<http://info.spsl.nsc.ru>), ГПИБ (<http://www.shpl.ru/>), ГПНТБ России (<http://www.gpntb.ru/>), ИНИОН РАН (<http://www.inion.ru/>), ЦНСХБ (<http://www.cnsxb.ru/>), НБ МГУ (<http://www.msu.su/>), ряда региональных и вузовских библиотек. Поиск можно провести и в электронном каталоге РГБ (<http://www.rsl.ru/>), но лишь по поступ-лениям последних лет.

Многие библиотеки публикуют на своих web-страницах текстовые файлы со списками библиографических указателей различных типов, дают перечни предоставляемых продуктов и услуг, рекламируя собственные информационные ресурсы, издания, которые они могут поставлять в традиционных режимах, открывают так называемые “электронные магазины”.

Полнотекстовых баз данных электронных журналов, представляющих интерес не только для библиотечкарей, на web-страницах библиотек практически нет. Чаше всего это труды сотруаников данной библиотеки и периодические издания по библиотечной тематике. Лишь ВИНТИ, ИНИОН и ВНТИЦ публикуют полные версии некоторых из издаваемых ими научных журналов, но эти организации и не являются чисто библиотечными учреждениями.

С другой стороны, работы по оцифровке фонда источников начинаются сейчас во многих библиотеках. Созаается хранилище электронных текстов источников в ИНИОН. В ГПНТБ СО РАН созана полнотекстовая БД прижизненных изааний А.С.Пушкина, завершается работа над электронной коллекцией аналитических обзоров по экологии. Не слишком большие аостижения библиотек в этой области объясняются не только отсутствием у них аостаточной материально-технической базы, но прежде всего проблемой соблюдения авторских прав. В обычной практике библиотеки редко принимают во внимание эту проблему, например, ксерокопируя источники по заказам своих читателей. То же копирование, только в электронном виде, не говоря уже о созании хранилищ электронных копий книг и журналов из библиотечных фондов и предоставление аостапа к ним через Интернет, является нарушением существующих законоаательных актов в области авторского права, весьма, наао сказать, несогласованных. Именно поэтому библиотеки сейчас выставляють в Интернете главным образом собственную продукцию, а она, за исключением библиографических указателей и списков литературы, ориентирована только на библиотечкарей.

Технологию электронной аоставки аокументов (ЭАА) освоили в экспериментальном режиме все крупнейшие библиотеки страны. ЭАА обеспечивает аауленный аостап к источникам, хранящимся в традиционном виде. В упрощенном виде здесь аействует следующая схема: читатель обрааается в свою библиотеку с просьбой заказать копию статьи из журнала, который она не выписывает. Заказ по электронной почте или через Интернет направляется в ту библиотеку, которая может его выполнить. Источник сканируется, и электронная копия пересылается библиотеке-аосреднику. Читателю выдается либо распечатка, либо файл на дискете.

Некоторые библиотеки (ИНИОН, РГБ, ГПИБ) имеют интер-фейс, позволяющий аауленному аользователю наарямую оформить заказ у исполнителя, минуя аосредника – местную библиотеку.

Несмотря на весьма простую и не очень затратную технологию, не так много библиотек перешли на промышленный режим ЭАА. К ним можно отнести (помимо уже упоминавшихся ИНИОН, РГБ и ГПИБ) ВИНТИ, ГПНТБ СО РАН, ЦНСХБ (1). Эти библиотеки, а также БАН, БЕН, ВПТБ, ВНТИЦ и некоторые другие сейчас объединяются в Ассоциацию российских служб ЭАА, целью которой является взаимодействие в обеспечении своих читателей и аауленных аользователей электронными копиями первоисточников из библиотечных фондов.

На первом этапе Ассоциация будет работать по следующей схеме. Любой читатель может обратиться за электронной копией в любую библиотеку, входящую в Ассоциацию. Заказ по электронной почте методом общей рассылки будет нааавлен сразу всем библиотекам – аленам Ассоциации. Библиотеки определяют, кто из них может выполнить заказ, и сообщают отправителю. Затем изготавливается копия и передается библиотеке-аказчику, которая в свою очередь передает ее читателю. Таким образом, читатели смогут через интерфейс ЭАА на сервере РГБ или ИНИОН заказать копию статьи, которая будет аделана, например, в ГПНТБ СО РАН в Новосибирске или в БАН в Санкт-Петербурге.

Несмотря на появление электронной аоставки аокументов, библиотеки продолжают совершенствовать межбиблиотечный абонемент. Развивается так называемый электронный МБА. Так, на протяжении двух лет стабильно работает встроенный в электронные каталоги БЕН, ГПНТБ СО РАН и др. модуль заказа по МБА через Интернет (2, 3).

Оако если мы внимательно посмотрим на описанные выше нааавления аеятельности библиотек в Интернете, то можем заметить, что это всего лишь новые технологические решения *традиционных* библиотечных процессов. И предоставление аауленного аостапа к собственным электронным каталогам, библиографическим, фактографическим и даже полнотекстовым базам аанных, и заказ, и выполнение услуг по каналам связи, и МБА – в принципе все это имело место и до появления Интернета. С

развитием техники и телекоммуникаций эти направления традиционной библиотечной деятельности лишь поменяли форму своей реализации.

### **Работа библиотек с ресурсами Интернета**

Решение второй группы задач – предмет активных дискуссий библиотекарей-практиков и библиотекарей-теоретиков. С ней связано также появление термина “Internet-librarian”. Помимо очевидного направления работы – поиска информации (библиографической, фактографической, полнотекстовой) в Интернете, выполнения справок по заказу удаленных пользователей и своих читателей, – библиотеки начали каталогизацию ресурсов Интернета. Эта работа первоначально развивалась (и продолжается до сих пор) как публикация на web-страницах аннотированных или неаннотированных списков ссылок на различные информационные ресурсы. Не имея возможности приобрести и сохранить саму информацию, библиотеки таким образом пытаются хотя бы обеспечить своим пользователям удобную навигацию в сети Интернет.

Однако накопление библиотеками таких ссылок и особенно поддержание их в актуальном состоянии весьма неэффективно. По опыту работы ряда библиотек средний срок нахождения информации по конкретному адресу – полгода (4). Правда, сейчас появились программы-роботы, позволяющие отслеживать ссылки, удерживать их в актуальном состоянии (см., например, (5)). Кроме того, как правило, ссылки создаются на весь информационный ресурс или его крупные части, а этого явно недостаточно. Если мы делаем ссылку, скажем, на электронное хранилище полных текстов журналов или книг, то что мы сообщаем пользователю? Информацию о том, что существует такой ресурс, не более. Это все равно, что сказать: “Ступайте в ближайшую библиотеку, там есть литература по интересующей вас теме”. Допустим, мы не просто сообщаем пользователю о существовании таких электронных коллекций, но и публикуем у себя перечень журналов и книг, которые они содержат. Это уже лучше, но все же мы опять отсылаем пользователя к этой электронной коллекции, предоставляя ему возможность самому отыскать конкретные публикации<sup>78</sup>. Ниже мы еще вернемся к этому вопросу.

Что же принципиально нового может появиться и появляется в работе библиотек с использованием и развитием сети Интернет? Возможно, в скором времени станет практикой архивирование самих информационных ресурсов Интернета. Прежде всего – это работа по описанию, учету, а возможно, и сохранению резервных копий этого нового вида информационных ресурсов – так, как это делается с обычными печатными изданиями в библиотеках и архивных хранилищах. Как мы уже говорили, это социальная задача библиотек, и сейчас она просто должна распространиться на новый вид ресурсов. Создатели этих ресурсов, так же как и издатели традиционных книг и журналов, в меньшей, чем библиотеки, степени ориентированы на сохранение электронных ресурсов для потомков. Пока, однако, этот вопрос даже не ставится, а его решение имеет массу сложностей – научных, организационных, правовых и технических.

### **Распределенные электронные библиотеки**

Еще одно направление работы – создание электронных библиотек – пошло по пути оцифровывания печатных источников и приобретения прав на электронные источники у издателей. При этом электронные фонды таких библиотек, как правило, автономны, замкнуты и расширяются лишь за счет тех же списков ссылок на другие ресурсы. Такие

---

<sup>78</sup> Форма представления информации на web-страницах библиотек носит несистемный характер с точки зрения распределения задач, аналогичного тому, которое сложилось в традиционной библиотечной среде и по линии Минкультуры, и РАН, не говоря уж о ГСНТИ. С одной стороны – имеет место массовое дублирование, начиная от web-адресов библиотек, причем не всегда полных и часто неточных, с другой – полное отсутствие систематизированных Интернет-ресурсов по тематике и проблемам, за информационную поддержку которых отвечает данная библиотека в традиционной системе. По нашей предварительной оценке, то, что библиотеки рекламируют и собирают на своих web-страницах – в 80% случаев имеет отношение к деятельности самих библиотек (данные о библиотеке, информация о конференциях, труды сотрудников, учеба, повышение квалификации и т.д.). Другими словами, web-страницы ориентированы в основном на библиотечное сообщество, а не на читателя, тем более – удаленного.

библиотеки состоят из разнородных информационных массивов — электронных каталогов, библиографических, фактографических, полнотекстовых баз данных, справочно-нормативных массивов и т.д. Пользователь, попадая в такую виртуальную автономную библиотеку, ограничен ее рамками и вынужден искать недостающую информацию (например, библиографические данные или полные тексты документов) за пределами такой коллекции, используя ссылки на некий совокупный ресурс.

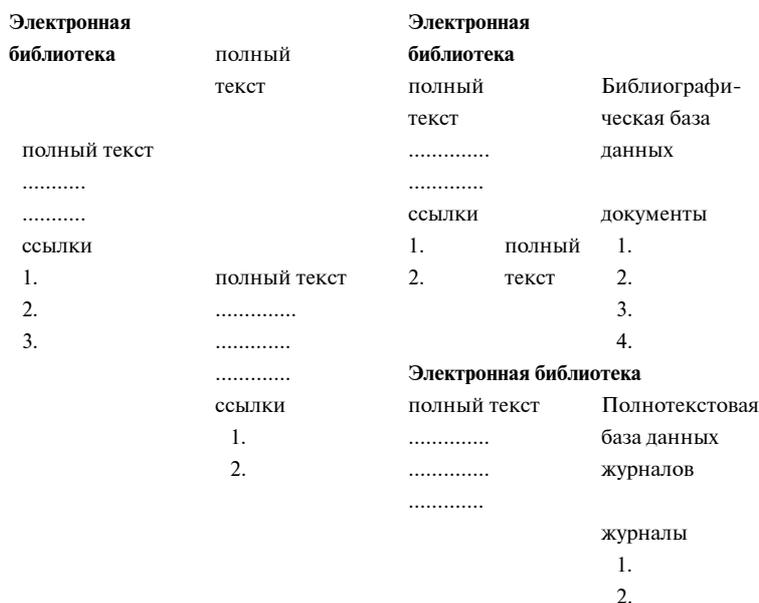
Конечно, создание подобного рода электронных библиотек является необходимым шагом, однако разработка таких локальных информационных продуктов, как нам кажется, — тупиковый путь развития.

Поэтому необходимо создание библиотек, обеспечивающих пользователей связями между элементами как своих, так и совокупных ресурсов, находящихся за их пределами.

В этой связи предлагаем рассмотреть следующую модель электронной библиотеки.

В отличие от традиционных библиотечных хранилищ, где фонд источников отделен от справочно-поискового аппарата в виде карточных каталогов, электронные библиотеки составляют единое целое, части которого связаны друг с другом, например, ссылками -библиографические описания с полными текстами. Помимо этих внутренних ссылок между частями отдельной электронной библиотеки создаются ссылки на внешние относительно данной коллекции ресурсы, причем не на отдельный информационный ресурс Интернета в целом, а на составляющие его единицы, например, конкретную статью из электронного журнала или другую публикацию, хранящуюся в полнотекстовой базе данных. Например, журнальные статьи из полнотекстовых баз данных с приставными списками литературы должны иметь ссылки, связывающие данную статью с полными текстами источников, которые в ней упоминаются. Эти ссылки или связи могут быть установлены между источниками, хранящимися как внутри одного электронного фонда, так и в разных коллекциях (рис. 1).

Подобный продукт может создаваться следующим образом. Любая научная публикация содержит некоторое число ссылок на использованную литературу. Вполне реален вариант, когда при подготовке научной работы автор заказывает и получает электронные копии, причем изготовитель этих копий обеспечивает их хранение; либо при подготовке рукописи используются статьи из электронных версий журналов и т.д. Такая публикация, представленная в Интернете (например, статьей в электронной версии журнала) и снабженная списком литературы, может быть связана через гипертекстовые ссылки с электронными копиями цитируемых публикаций, которые находятся в доступных через Интернет хранилищах, так же как и со статьями из электронных журналов. Технологической основой к созданию такого информационного продукта может служить, в частности, технология электронной доставки документов. Главное то, что эта технология позволяет объединить разобщенные информационные ресурсы и преодолеть пространственные барьеры между ними.



.....

**Рис. 1. Связи между элементами распределенной  
электронной библиотеки**

Такую библиотеку можно назвать распределенной электронной библиотекой. *Распределенная электронная библиотека – это совокупность логически и технологически связанных между собой документов из различных электронных коллекций и отдельных информационных ресурсов, доступных через Интернет.*

Что следует включать в состав фонда такой библиотеки, какие связи с другими ресурсами она должна обеспечивать?

Нам представляется, что распределенная электронная библиотека может характеризоваться:

- принципами отбора документов (комплектования);
- способами создания связей между элементами;
- методами формирования электронного фонда.

Что касается принципов отбора документов, то из опыта работы с информационными ресурсами Интернета следует, что части фонда распределенной электронной библиотеки должны быть элементами профессионально создаваемых и постоянно поддерживаемых информационных ресурсов. Точно так же, как традиционные библиотеки не включают в свои фонды некоторые виды печатной продукции (что определяется, конечно, статусом библиотеки), – так и распределенные электронные библиотеки должны охватывать не все информационные ресурсы, а лишь те из них, которые входят в область интересов их пользователей.

Основной проблемой при создании подобной библиотеки является генерация гипертекстовых ссылок на внешние ресурсы и их фрагменты – связей между распределенными в пространстве элементами электронных библиотек. Сложность заключается в том, что пока не существует уникального идентификационного номера электронной публикации в Интернете наподобие ISSN или ISBN, хотя разработки в этой области ведутся как издателями, так и библиотечным сообществом.

Определенные надежды здесь возлагаются на описания типа Dublin Core (6), хотя, как нам кажется, это совершенно необоснованно, и вот почему. Во-первых, пока эти описания не распространяются на динамически формируемые документы из баз данных и, во-вторых, сама процедура описания отдается на откуп создателям Интернет-ресурсов. Вспомним, чем закончились попытки обязать авторов и издателей книг делать библиографические описания на свои произведения. Как правило, эти описания непригодны для использования в библиотеках, потому что создаются непрофессионалами в этой области. С Dublin Core дело обстоит еще сложнее, поскольку предполагается, что авторы будут сами индексировать свои электронные издания. С результатами подобного “самоиндексирования” может ознакомиться каждый, кто хоть раз пытался найти ссылки в поисковых системах типа Yahoo или Rambler. Поэтому, как нам кажется, до тех пор пока не разработаны надежные средства описания информационных ресурсов Интернета, следует сосредоточить усилия на разработке систем автоматического обнаружения и описания этих ресурсов, средств эвристического поиска.

Очевидно, что сопряжение элементов распределенных электронных библиотек (процедура установления ссылок между их частями) должно осуществляться автоматически, поскольку, во-первых, только так могут быть учтены все связи, и, во-вторых, это должно позволить поддерживать их в актуальном состоянии – так как элементы электронной библиотеки меняют свое местоположение, добавляются новые документы и т.д. То же относится и к связям, которые устанавливаются между документами локального электронного фонда. Такие разработки сейчас ведутся рядом российских библиотек.

Методы формирования распределенных электронных библиотек могут быть различными. Условно их можно назвать прямым и косвенным. Прямой метод заключается в том, что части таких библиотек формируются как результат совместной деятельности, т.е. целенаправленно создается распределенный электронный фонд. Так, при подготовке некоторых изданий, составляющих основу специализированной электронной библиотеки, могут быть оцифрованы источники, необходимые для их написания. Эти источники входят затем в состав электронного фонда копий публикаций другой электронной библиотеки и могут быть связаны ссылками с текстами первой коллекции. Такой метод формирования

распределенного фонда наиболее продуктивен, поскольку повышается качество отбора документов в фонд библиотеки.

Косвенный метод заключается в том, что автоматически устанавливаются отсылки от документов фонда одной электронной библиотеки и их структурированных частей к документам других электронных библиотек. Для этого должны быть разработаны системы справочников, описывающих структуры этих библиотек и эвристические программы обнаружения электронных информационных ресурсов.

Отметим, что в подобных библиотеках может существовать два вида ссылок – внутренние, связывающие части одной коллекции между собой, и внешние, которые устанавливаются между элементами различных коллекций.

Рассмотрим теперь конкретный пример создания подобной библиотеки, базирующейся на использовании указанных методов. В качестве исходных электронных библиотек взяты:

- Электронная библиотека ИНИОН (<http://www.inion.ru>). Она состоит из библиографических баз данных, содержащих описания книг и статей из журналов, поступивших в библиотеку ИНИОН, а также хранилища электронных копий источников, отсканированных по заказам пользователей или в рамках специальных проектов;

- Научная электронная библиотека (<http://library.intra.ru>), созданная по инициативе Российского фонда фундаментальных исследований. Она содержит полные тексты статей свыше 350 зарубежных и отечественных электронных журналов;

- Электронная коллекция ГПНТБ СО РАН по экологии <http://info.spsl.nsc.ru/win/ecol/index.html>, имеющая полные тексты аналитических обзоров, представленные в Интернете.

При подготовке обзоров ГПНТБ СО РАН по экологии проводились поиски информации по различным библиографическим базам данных, в том числе базам ИНИОН, а затем заказывались копии в службе электронной доставки документов этой организации. В некоторых случаях в электронном хранилище ИНИОН оказывались полные тексты статей, которые заведомо процитированы в обзорах или указаны в списке литературы. Таким образом, гипертекстовая ссылка от библиографического описания источника в списке использованной литературы к документу из электронного полнотекстового хранилища ИНИОН позволяет увязать документы из разных частей распределенной библиотеки.

Электронные тексты первоисточников, хранящиеся в ИНИОН, равно как и другие элементы двух электронных библиотек, в свою очередь имеют гипертекстовые отсылки, например, к другим работам автора обзора в этом же хранилище электронных копий, или – опять через список литературы – к другим хранилищам и базам данных.

Для подготовки обзоров также используются статьи из коллекции электронных журналов Научной электронной библиотеки. Через разрабатываемую сейчас систему справочников эти элементы распределенной электронной библиотеки также должны быть связаны друг с другом. Таким образом, формируются единые тематические или проблемно ориентированные информационные поля в рамках распределенной библиотеки. В нашем случае – по различным аспектам экологии и охраны окружающей среды. Сейчас ведутся работы по реализации этой технологии на примере создания информационных полей в области экономических и правовых аспектов охраны окружающей среды.

Полагаем, что сказанное выше открывает путь к созданию новых электронных библиотек на базе формально распределенных в пространстве информационных ресурсов – библиотек, создание которых принципиально невозможно в традиционной информационной среде.

### Список литературы

1. Глухов В.А., Лаврик О.Л. Электронная доставка документов. – М.: ИНИОН РАН, 1999. – 132 с.
2. Каленов Н.Е, Власова С.А. Каталог книг и продолжающихся изданий ЦБС БЕН РАН в Интернете // Современные технологии в информационно-библиотечном обеспечении научных исследований: материалы 8-го научно-практического семинара (Таруса, 13-18 сентября 1999 г.). – М., 1999. – С. 100-102.; Соловьева С.А. Система баз данных на основе электронного журнального каталога БЕН РАН, доступная через Интернет // Там же. – С. 97-99.

3. Баженов С.Р., Баженов И.С., Мазов Н.А., Малицкий Н.А. Система обеспечения контроля и учета доступа из Интернета к распределенным библиографическим базам данных // Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества. Материалы конф. Судак. Респ. Крым, Украина, 5-13 июня 1999 г. – М., 1999. – С. 226-228; Баженов С.Р., Новоженин А.Ф. Разработка системы решения информационно-библиотечных задач в ГПНТБ СО РАН с использованием Интернета // Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества. Материалы конф. Судак. Респ. Крым, Украина, 6-14 июня 1998 г. – М., 1998. – С. 265-267.
4. Кон Л.Ф., Дашенко В.В. Интернет в Новосибирской государственной областной библиотеке // Тезисы докладов Междунар. конф. “Интернет, общество, личность”, Санкт-Петербург, 1-5 февраля 1999 г. – СПб, 1998. – С. 104-106.
5. Паринов С. Активный информационный робот как сетевой агент исследователя на примере онлайн-ресурсов по экономике RePEc/Ru/RuPEc // Электронные библиотеки: перспективные методы и технологии, электронные коллекции: Тр. первой Всероссийской научной конференции, С.-Петербург, 19-22 октября 1999 г. – СПб, 1999. – С. 104-111.
6. Шварцман М.Е. К вопросу каталогизации ресурсов Интернета // Мир библиографии. – 1998. – N5; Шварцман М.Е. Использование метаданных для каталогизации российских ресурсов Интернета // Электронные библиотеки. – 1998. – Т. 1, Вып. 2.