



- ▶ Журнал ЭБ
- ▶ О журнале
- ▶ Редакционная коллегия и редакция
- ▶ Что нового?
- ▶ 2014 год
- ▶ 2013 год
- ▶ 2012 год
- ▶ 2011 год
- ▶ 2010 год
- ▶ 2009 год
- ▶ 2008 год
- ▶ 2007 год
- ▶ 2006 год
- ▶ 2005 год
- ▶ 2004 год
- ▶ 2003 год
- ▶ 2002 год
- ▶ 2001 год
- ▶ 2000 год
- ▶ 1999 год
- ▶ 1998 год

▶ ENGLISH

Электронные библиотеки -1999 - Том 2 - Выпуск 2

Информация + технологии: опыт сотрудничества Библиотеки Российской академии наук и Института высокопроизводительных вычислений и баз данных

С.П. Голенский, Е.Г. Пустельникова
Институт высокопроизводительных вычислений и баз данных

В.П. Захаров, А.Ц. Масевич
Библиотека Российской академии наук

С течением времени информация становится одним из наиболее значимых стратегических ресурсов, оказывающих решающее воздействие на изменения в социальной структуре общества, экономики, развитии институтов демократизации общества.

Международный проект "Цифровые (электронные) библиотеки" (the Digital Library Initiative, DLI) появился в результате возрастающей потребности в системах и службах, способных обеспечивать различные категории пользователей доступом к необходимым им большим и географически распределенным источникам информации.

Целью проекта стало дальнейшее развитие средств и методов, предназначенных для сбора, хранения, организации данных в электронной форме и обеспечения доступа к ним посредством использования компьютерных сетей с целью поиска, получения и обработки необходимой информации. Проводимые в рамках проекта исследования также сфокусированы на социальных, экономических и культурных аспектах развития и использования электронных ресурсов.

Несмотря на сложное экономическое и политическое положение в России, необходимость ее полноценного участия в существовании и развитии всемирного информационного пространства становится все более и более очевидной.

Развитие национальной программы "Российские электронные библиотеки" ограничивается текущим экономическим и технологическим положением страны. Одна из целей программы может быть сформулирована как предотвращение дальнейшей потери ранее имевшегося статуса одного из мировых центров научной, технологической, культурной и гуманитарной активности.

Накопленные к настоящему моменту и постоянно пополняющиеся обширные информационные массивы могут (и должны) составить основу будущих систем хранения и представления информации пользователю. А библиотечные фонды представляют собой, вероятно, наиболее полную, логически стройную, завершенную и стабильную составляющую в огромном информационном пространстве России.

Тенденции создания и развития электронных библиотек, потребность в хранении, обработке и передаче больших объемов данных приводят к необходимости использования в этой области передовых информационных технологий, включающих в себя как новейшее программное обеспечение, так и высокопроизводительные компьютерные системы. И несмотря на то, что за последнее время произошли существенные изменения в оснащении библиотечных организаций вычислительной техникой, средствами оцифровки и передачи информации, каналами выходов в телекоммуникационные сети (в том числе в Интернет), широкое использование сложных, дорогостоящих, требующих значительных вычислительных и человеческих ресурсов технологий затруднено в условиях недостаточного финансирования российских библиотек. В этих условиях большое значение приобретает сотрудничество между коллективами, достигшими определенных успехов в конкретной области деятельности.

В рамках такого сотрудничества в настоящее время Библиотекой Российской академии наук (БАН) и Институтом высокопроизводительных вычислений и баз данных (ИВВиБД) реализуется программа совместной деятельности в области интеграции информационных массивов библиотеки и ресурсов компьютерного центра, что позволяет применять современные технологии к большим накопленным объемам уникальных данных.

Библиотека Российской академии наук - первая государственная библиотека в России, основанная Петром I в 1714 году и получившая статус академической в 1725 году. Библиотека имеет уникальные фонды, привлекающие внимание исследователей во всем мире, и является административным и

методическим центром сети академических библиотек.

Институт высокопроизводительных вычислений и баз данных Министерства науки и технологий развивает новые технологии преимущественно в области использования суперкомпьютерной техники и осуществляет информационную и вычислительную поддержку государственных учреждений, исследовательских и образовательных организаций соответственно уровню национального центра. Это подразумевает, в том числе, подготовку и концентрацию специалистов в области архитектуры и использования супер-ЭВМ для решения актуальных задач современности. Целью деятельности Института как головного в своей области является внедрение современных суперкомпьютерных технологий путем формирования устойчиво развивающегося контингента потребителей, использующих алгоритмы, модели, программные средства и вычислительные мощности ИВВиБД. Институт обладает мощным аппаратно-программным комплексом, на основе которого может быть создан высокопроизводительный сервер баз данных.

На первом этапе совместной деятельности БАН и ИВВиБД для публикации данных в сети Интернет применяется "Типовое решение для библиотечных Internet-серверов WebISIS/Linux, подготовленное Государственной публичной научно-технической библиотекой России в рамках проекта LIBWEB. Этот путь был выбран, поскольку в качестве ядра открытой информационной системы используется CDS/ISIS на платформе Linux, а специалистами БАН уже в течение ряда лет успешно эксплуатируется свободно распространяемая система управления библиографическими базами данных CDS/ISIS (под операционной системой MS-DOS), во внутренних форматах которой и представлена рассматриваемая информация. Так что обмен данными между родственными системами, пусть и на различных платформах, не должен был составить значительных сложностей. Доступ пользователей к библиографическим ресурсам осуществляется по протоколу HTTP средствами WWW-шлюза Web-ISIS и стандартных Web-browser'ов. Компоненты шлюза предназначены для организации доступа пользователей Интернет к библиографическим базам данных, поддерживаемых в среде CDS/ISIS.

На настоящий момент, пока еще в экспериментальном режиме, в Интернет (на WWW-страницах БАН сервера ИВВиБД - www.csa.ru/BAN) опубликованы следующие ресурсы БАН:

- Поступления иностранных монографий в БАН с 1 января 1997 года
- Бюллетень новых поступлений в кабинет библиотековедения БАН
- Ежегодник "Библиография изданий академии наук"
- Материалы к библиографическому изданию "Советская/российская археологическая литература", выпускаемому библиотекой института истории материальной культуры
- База данных "Bryophyta" - библиография литературы по мхам, которая ведется библиотекой ботанического института им. Комарова.

В самом ближайшем будущем планируется публикация материалов библиотек физико-технического института им. Иоффе и института озероведения.

Доступ к перечисленным материалам является завершением первой фазы совместной деятельности. Участники проекта намерены и далее продолжать сотрудничество, приступив к решению следующих задач, которые теперь уже можно сформулировать, систематизировать, разбить на несколько групп и решать по мере сил и возможностей.

Технологические задачи

На следующем этапе сотрудничества участникам предстоит разработать технологию адекватного ситуации обновления опубликованных в Интернет баз данных, поскольку все доступные сегодня материалы постоянно обновляются; библиотека создает в неделю до тысячи единиц библиографической информации разного характера. Необходимо определить технические средства, оптимальные порядок и частоту пополнения баз данных на сервере. Учитывая слабую техническую оснащенность БАН, эта задача вряд ли будет простой.

Доступ к данным через Интернет является формой их публикации. Многие материалы, установленные в ходе первой - экспериментальной - фазы проекта, носят в значительной степени предварительный характер, они нуждаются в дополнительном редактировании. По этой причине мы считаем необходимым разработать в будущем издательский цикл, включающий планирование, создание, редактирование, рецензирование и порядок выпуска электронных изданий.

Техническое и программное обеспечение

Для оперативной реализации технологических разработок необходимо применение современных средств передачи электронной информации, а также, возможно, разработка или адаптация существующих программных продуктов, обеспечивающих учет и регулярность обновления ресурсов на сервере.

Как уже сказано выше, использован программный продукт, разработанный в ГПНТБ России для библиотек - участниц проекта LIBWEB в качестве типового решения представления данных пакета CDS/ISIS в Интернет. Это программное средство, полноценно выполняя свои функции, тем не менее, не вполне удовлетворяет нашим потребностям. На представляемых данных для обслуживания

пользователей используются несколько специфических программных модулей на языке ISIS Pascal, созданных в БАН, которые, естественно, не имеют адекватного представления средствами WWW-шлюза. Кроме того, поисковые средства этого продукта представляются недостаточными. Поэтому мы полагаем необходимой модификацию разработки ГПНТБ (а может быть, и создание своего собственного программного средства) одновременно с другой работой, запланированной на 1999 год – созданием пользовательского интерфейса для информационных систем БАН.

Организационные задачи

Предстоит также решить непростую и неоднозначную при существующей практике проблему – о соблюдении авторской ответственности и авторских прав. Кроме того, возникает вопрос об условиях распространения изданий. Доступ к некоторым из созданных совместно ресурсов может осуществляться на коммерческой основе. В этом случае неизбежно встают вопросы о собственности на информацию и на совместно созданные программные продукты. Отметим, впрочем, что на сегодняшний день ни та, ни другая из участвующих в реализации проекта организаций не ставят перед собой коммерческих целей.

Кроме перечисленных задач следует, вероятно, коснуться более отдаленной перспективы, и прежде всего – создания библиотеки компьютерных копий особо ценных документов. По величине и ценности фондов (свыше 19 млн. изданий, рукописей, карт, гравюр и других документов) БАН входит в число шести великих библиотек мира. В ее фондах хранятся ценнейшие документы. Среди них русские летописи 15-го века – Ипатьевская и Радзивилловская, три тома Лицевого летописного свода Ивана Грозного, уникальные собрания русских рукописных карт, гравюр петровского времени, книги с пометами, сделанными Петром Великим, Галилео Галилеем, М.В. Ломоносовым, Леонардом Эйлером, Бенджамин Фрэнклином и много других печатных и рукописных ценностей.

Вполне естественно, что доступ к этим материалам ограничен. Создание цифровых копий особо ценных изданий и доступ к ним через Интернет – следующее возможное направление нашего сотрудничества.

Создание компьютерной копии – это не только доступ многих пользователей к ценному документу, но еще и гарантия его сохранности. Помимо этого средства, позволяющие увеличить фрагмент изображения, будут, вероятно, полезны специалистам, изучающим рукописи и старопечатные книги.

Для реализации столь далеко идущих планов в перспективе применение высокопроизводительной техники параллельной архитектуры, программного обеспечения ведущих производителей и высокоскоростных каналов передачи данных обеспечит увеличение производительности информационных систем и скорости передачи данных, возможность создания полнотекстовых и мультимедийных баз большого объема, многофункциональный поиск по каталогам и базам данных. На данном этапе проводится предварительный анализ и начальное проектирование будущей распределенной базы данных, освоение существующих технологий и изучение перспективных направлений развития как компьютерной, так и библиографической областей, обобщение уже полученного, пусть небольшого, но очень важного опыта взаимного сотрудничества. При физической реализации будущей информационной системы предполагается использовать реляционную модель данных, продукт корпорации **IBM DB2 UDB V5.0** в качестве ядра проектируемой информационной системы, платформу Sun Solaris и обеспечение пользовательского доступа через Интернет средствами стандартных Web-browser'ов.

Выбор реляционного подхода обусловлен в основном двумя факторами. Во-первых, база данных представляется на внешнем, не зависящем от структуры компьютера уровне как совокупность повседневно встречающихся в человеческой практике двумерных таблиц, в которых поиск и обработка информации не зависят от организации хранения данных в памяти компьютера. Это представление значительно упрощает взаимодействие пользователя с банком данных и значительно повышает производительность его труда. Во-вторых, реляционная база данных с математической точки зрения – это конечный набор конечных отношений различной арности между заранее определенными множествами элементарных данных. Другими словами, реляционная база данных – это конечная модель в смысле математической логики. Над отношениями модели можно осуществлять различные алгебраические операции. Нормализация отношений обеспечивает целостность и повышает эффективность использования базы данных в целом.

Анализ существующих на сегодняшний день решений и их оценка с точки зрения соответствия поставленным задачам и затрат, связанных с использованием этих решений, обусловили выбор программной и аппаратной платформ – использование продукта корпорации IBM DB2 UDB V5.0 в качестве ядра проектируемой информационной системы и платформы Sun, Solaris. Дополнительный положительный момент такого решения – возможность приобретения всех необходимых компонент (аппаратная платформа и СУБД) у одного поставщика.

Конечной целью построения любой информационной системы является получение пользователем некоторой информации. В этой связи немаловажную роль играет форма подачи этой информации. Для того, чтобы пользователь мог легче воспринимать большие объемы информации, разработано множество методов ее представления. Все эти разработки учитываются при проектировании пользовательского интерфейса, то есть той части программного комплекса, представляющего

информационную систему, которая непосредственно будет взаимодействовать с пользователем. В таком случае представляется наиболее оптимальным использовать такой пользовательский интерфейс, который может быть видоизменен каждым конкретным пользователем для себя. Предусмотренная возможность формирования пользовательского интерфейса облегчает усвоение большого объема разнообразной информации, предоставляемой информационной системой.

При разработке информационных систем представляет интерес также возможность получения интересующей информации посредством использования глобальных электронных сетей, например Internet. При работе в сети Internet наиболее удобным является использование стандартных программ для навигации по ресурсам сети. Например, таковыми являются Internet Explorer и Netscape Navigator. Таким образом, с одной стороны, отпадает необходимость для разработчика писать множество версий одного и того же продукта для различных операционных систем, а для пользователя разыскивать и настраивать соответствующую версию при переходе к другой системе; с другой стороны - обеспечивается общность пользовательского интерфейса.

Сейчас уже можно с большой долей вероятности предсказать неминуемое нарастание процесса расширения компьютерно-библиотечной сети России, увеличение числа онлайн-центров, сотрудничающих и конкурирующих между собой. Будучи явлением несомненно положительным, компьютеризация вызывает появление новых проблем. Перед работниками библиотек встает задача описания (а отчасти - и предсказания) требуемой функциональности будущих систем, способной полностью обеспечивать требования конечных пользователей Интернет, возможно, отличные от запросов посетителей традиционных библиотек. Специалистам компьютерных центров, в свою очередь, предстоит решать проблемы представления библиографической информации во Всемирной Паутине, обработки и передачи больших объемов данных по несовершенным еще каналам связи.

© *Голенский С.П., Захаров В.П., Масевич А.Ц., Пустельникова Е.Г., 1999*

Последнее обновление страницы было произведено: 2003-12-09

Все предложения и пожелания по содержанию и структуре портала направляйте по адресу rdlp@iis.ru

