

**Программное обеспечение, позволяющее авторам создавать
полнотекстовые электронные документы и производить
их автоматическую каталогизацию**

**Software Which Enables Authors to Produce
Full-text Digital Documents and Carry Out
their Automatic Cataloguing**

**Програмне забезпечення, що дозволяє авторам створювати
повнотекстові електронні документи та виконувати
їх автоматичну каталогізацію**

Л. Г. Еремеев, А. В. Кузнецов

*Омский государственный университет, Институт математики
и информационных технологий, Омск, Россия*

Leonid Eremeev and Alexander Kuznetsov

*Omsk State University, Institute of Mathematic and Information Technologies,
Omsk, Russia*

Л. Г. Єремєєв, О. В. Кузнецов

*Омський державний університет, Інститут математики
та інформаційних технологій, Омськ, Росія*

Освещается комплексный подход к созданию электронных полнотекстовых библиотечных материалов с их автоматической каталогизацией, размещению их описаний в библиотечных электронных каталогах, и осуществлению доступа к ним с использованием возможностей протокола Z39.50.

The authors describe an integrated approach allowing for producing digital full-text library documents and carrying out their automatic cataloguing, placing their entries in library electronic catalogues, and providing access to them by means of Z39.50.

Висвітлюється комплексний підхід до створення електронних повнотекстових бібліотечних матеріалів із їх автоматичною каталогізацією, розміщенням їх описів у бібліотечних електронних каталогах, і здійснення доступу до них із використанням можливостей протоколу Z39.50.

На сегодняшний день в России существует достаточно мало электронных каталогов библиотек, в которых помимо электронных описаний документов есть и ссылки на полные тексты этих документов, доступных через сеть Интернет. Приятным исключением является электронные каталоги Российской государственной библиотеки, среди которых есть каталоги, содержащие ссылки на полнотекстовые материалы. РГБ предоставляет читателям возможность доступа уже к достаточно большому количеству полнотекстовых материалов.

Однако этот пример является на сегодняшний день скорее исключением, чем правилом.

В то же время можно констатировать тот факт, что люди привыкли искать и использовать полнотекстовые документы в электронном виде, и все меньше обращаются в библиотеки за их бумажными вариантами. Сложилась ситуация, когда библиотеки уже все меньше и меньше удовлетворяют информационные потребности читателей. Т. е. библиотеки не занимаются созданием полнотекстовых электронных коллекций.

Конечно, существует ряд причин, которые обуславливают такое положение дел, например следующие:

- отсутствие финансирования данного направления,
- отсутствие данных работ в производственных планах библиотек,
- проблемы, связанные с авторскими и смежными правами, в том числе правами на публикацию документов.

Но в тоже время достаточно большое количество авторов самостоятельно осуществляют электронные публикации своих работ, размещая эти работы на тех или иных сайтах, порталах, собственных электронных библиотеках.

Однако в абсолютном большинстве случаев для поиска этих электронных материалов приходится использовать не библиотечные технологии.

Для получения доступа к полным текстам, пользователям необходимо либо изучать структуру каждого из полнотекстовых сайтов или порталов в отдельности, либо пользоваться поисковыми системами общего назначения, которые имеют ряд недостатков, связанных именно со своей спецификой:

- ограниченность поисковых возможностей (например, в них нет возможности искать документы конкретного автора),
- область поиска гораздо шире множества полнотекстовых документов.
- Таким образом, решалось две задачи:
- помочь авторам публиковать свои полнотекстовые материалы, а их электронные описания размещать в библиотечных каталогах, используя для этого международный стандарт – MARC-формат;
- дать возможность читателям находить их с использованием библиотечных технологий поиска.

Для решения этих задач в Институте математики и информационных технологий Омского государственного университета было разработано инструментальное программное обеспечение, позволяющее создавать полнотекстовые материалы и производить автоматическую каталогизацию (создание записи электронного каталога) электронного документа (статьи, учебно-методического пособия, монографии) без обращения автора к профессиональным программистам и к библиотечным специалистам (каталогизаторам).

Конечно, создаваемые таким образом записи электронного каталога уступают по качеству записям, которые могут создать профессиональные каталогизаторы, но все основные данные, по которым читатель привык искать необходимый материал в библиотечном каталоге, эти записи содержат.

В состав данного программного обеспечения входят, в частности, два АРМа: АРМ «Автор» и АРМ «Администратор».

При разработке АРМа «Автор» преследовались следующие цели: дать автору легкий в освоении инструмент, который не требует специальных знаний и позволяет создавать материал без привлечения к этой работе профессионального программиста.

Созданный АРМ «Автор» удовлетворяет этим требованиям. Автор может создавать свое произведение, используя различные текстовые редакторы (например, Word), может вставлять в текст изображения, формулы, гиперссылки и т. п. Содержимое создаваемого документа может быть сохранено в большинстве из наиболее распространенных форматов данных: doc, txt, pdf, djv. С помощью этого АРМа можно не только создавать новые материалы сразу в электронном виде, но и легко создавать электронные версии материалов, опубликованных ранее с применением традиционных бумажных технологий.

После того, как автор заканчивает работу над своим произведением, ему предлагается заполнить естественно воспринимаемые им поля: название работы, ФИО автора (авторов), аннотацию, ключевые слова, год создания. Содержимое этих полей используется для создания карточки электронного каталога и автоматической каталогизации. Кроме этого, автору предлагается выбрать из списка существующих электронных каталогов тот, который, по его мнению, лучше подходит к создаваемому произведению.

Такой подход существенно ускоряет внесение информации о работе автора в электронный библиотечный каталог. А в тех случаях, когда по каким-то причинам в библиотеке не налажена работа по созданию полнотекстовых библиотечных электронных баз данных (а может и вообще отсутствовать библиотека со своей АБИС), этот подход позволяет автору осуществить электронную публикацию с использованием библиотечных технологий.

Практика показала, что для освоения этого АРМа большинству авторов достаточно одного-двух дней.

АРМ «Администратор» создан для управления работой по сопровождению базы данных, содержащей полнотекстовые материалы, сопровождению электронных каталогов, учетных записей авторов, статистического отслеживания содержимого каталогов.

Администратор фактически является управляющим электронной библиотекой, содержащей полнотекстовые материалы. Он создает электронные каталоги, создает учетные записи авторов, анализирует статистику.

Реальное создание записей и размещение их в электронных каталогах происходит в тот момент, когда администратор решит произвести обновление этих каталогов. Это решение он принимает на основе анализа статистических данных, в которых, в частности, содержатся сведения о том, сколько новых работ готовы к публикации.

Созданное программное обеспечение относится к классу клиент-серверных систем. Для ее работы на стороне сервера необходимы следующие компоненты:

- операционная система,
- веб-сервер,
- СУБД,
- АРМ «Автор»,
- АРМ «Каталогизатор»,
- Z-сервер.

И автор и администратор могут использовать эту систему удаленно, находясь либо на работе, либо дома, либо в командировке. Необходимо только доступ в Интернет и наличие в своем компьютере какого-либо веб-браузера (например, Internet Explorer, Firefox, Opera)

Поиск по электронному каталогу может быть осуществлен с использованием любого Z-клиента. Для просмотра полных текстов документов необходим веб-браузер.

В настоящее время в Институте математики и информационных технологий ОмГУ такая система реализована с использованием следующих компонент:

- ОС Microsoft Windows XP,
- Веб-сервер Apache,
- СУБД Postgres,
- АРМ «Автор»,
- АРМ «Каталогизатор»,
- Z-сервер Zebra.

Адрес сервера: imit.omskreg.ru, порт 9999

С ее помощью в данный момент созданы и доступны электронные каталоги:

- ege – ЕГЭ по математике,
- ethno – полнотекстовые материалы по этнографии,
- ptd – полнотекстовые статьи и книги.

Данные каталоги включены в дерево библиотечных ресурсов Z-клиента LibNavigator, который (вместе с деревом ресурсов) можно загрузить с сайта www.libnavigator.ru.

Каждая запись в этих каталогах содержит ссылку на свободно доступный полнотекстовый материал.