

# О двух подходах к созданию электронной библиотеки научно-технической документации в формате TeX

© М.О. Коваленко, А.В. Цуриков, М.К. Чернышов

Воронежский государственный университет  
maxel@vsau.ru

## Аннотация

В данной статье рассматриваются два частных подхода к решению задачи создания электронной библиотеки научно-технической документации в формате TeX. Дается краткое описание структуры данных библиотеки, приводятся аргументы в пользу выбора используемых средств разработки.

Данная работа посвящена рассмотрению вопросов, связанных с разработкой структуры электронной библиотеки научно-технической документации, созданием интерфейса пользователя библиотеки, организацией процесса доступа к документации по запросу пользователя, а также реализации механизма отображения требуемых документов на компьютере пользователя электронной библиотеки.

В рамках решения поставленной задачи в качестве исходного формата документов выбран формат LaTeX, являющийся макрорасширением формата TeX, который позволяет гармонично совмещать текст с математическими формулами, таблицами, диаграммами, графическими объектами. Поскольку документы электронной библиотеки в формате TeX могут занимать различный объем памяти, использование реляционных баз данных для формирования структуры библиотеки представляется громоздким и не совсем логичным, в связи с чем предлагается рассмотреть два альтернативных подхода к решению данной задачи.

**В первом случае** в качестве основы при создании электронной библиотеки выбрано дерево каталогов файловой системы ext2 (ext3/ext4) семейства операционных систем Linux. При этом выбор операционной системы Linux в качестве платформы для создания данного программного продукта не случаен. Используемая в данном семействе операционных систем файловая система позволяет создавать жесткие и символические ссылки на файлы, что дает возможность, в свою очередь, существенно упростить структуру базы документов и, как следствие, организацию процесса поиска в библиотеке требуемых документов, не создавая для этого дополнительных громоздких деревьев каталогов.

При осуществлении доступа к библиотеке используется технология «клиент-сервер», при реализации которой пользователи получают доступ к документам библиотеки, расположенным на сервере, с удаленных компьютеров. Для осуществления механизма доступа к удаленному серверу с использованием технологии CGI-программирования применяется язык программирования PERL как мощное средство обработки текстов, взаимодействия с базами данных и обеспечения безопасности при работе с компьютерными сетями. Для работы с документами создано Web-приложение, позволяющее пользователю добавлять в библиотеку новые документы, осуществлять многоуровневые запросы для поиска документов, удовлетворяющих требуемым критериям, а также производить редактирование существующих документов. При этом отдельно решен вопрос о разграничении прав доступа для различных категорий пользователей с целью обеспечения безопасности хранения документации.

Поиск данных в библиотеке реализован следующим образом: имеются директории, имена которых соответствуют метаданным книг: название, год издания, издательство, авторы и так далее. В этих директориях хранятся символические ссылки на соответствующие книги, и при запросе пользователя запускается PERL скрипт, который ищет соответствующие ссылки в этих директориях.

Для получения доступа по компьютерной сети к документам электронной библиотеки, расположенным на сервере, и при разработке Web-интерфейса пользователя используется Web-сервер Apache. При помощи языка программирования PERL созданы элементы интерфейса пользователя, инструменты для взаимодействия пользователя с документами электронной библиотеки. Созданы скрипты для работы администратора, реализующие взаимосвязь между механизмом добавления или редактирования новых пользователей, механизмом авторизации, механизмом поиска данных в библиотеке по запросу пользователя и механизмом отображения найденного документа на компьютере пользователя. Процесс авторизации для доступа к электронной библиотеке является трехуровневым (администратор, зарегистрированный пользователь и гость). Для отображения найденных по запросу пользователя документов в окне Web-браузера использован конвертер LaTeX2HTML, служащий для преобразования TeX-документов в формат HTML.

---

Труды 13<sup>й</sup> Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» - RCDL'2011, Воронеж, Россия, 2011.

**Второй подход** к решению задачи создания библиотеки электронной документации в формате TeX основан на использовании сервера приложений IBM Lotus Domino Server.

Отличие от большинства систем управления базами данных, в которых организация данных имеет четко структурированную табличную форму, в системе IBM Lotus Domino единицей хранения информации является документ. Все системы, основанные на первом принципе, требуют описания структуры данных, а реализующие реляционную модель со свойствами поддержания ссылочной целостности также требуют описания взаимосвязей между объектами. В IBM Lotus Domino структура данных может меняться в процессе работы без опасений относительно искажения или потери этих данных.

Взаимодействие пользователя с документами в Lotus Domino производится посредством форм, в которых существует возможность изменения информации документа, причем на структуру документа влияет только информация, введенная пользователем посредством формы. Документ вначале создается, и только затем он предоставляется пользователю для возможности изменения его структуры, в отличие от систем табличного типа, где, прежде чем создать единицу хранения, система требует описания структуры данных, а только затем создается сама единица хранения.

Представленный ниже алгоритм описывает принцип работы построенной электронной библиотеки:

1. Пользователь вводит адрес Web-страницы электронной библиотеки, и Domino предлагает ему авторизоваться. Пользователь должен ввести имя и пароль, полученные во время регистрации.
2. На основании введенных данных сервер Domino предоставляет пользователю один из нескольких уровней доступа. Система предоставляет возможность авторам изменять и добавлять свои документы, а пользователям просматривать документы. Помимо этого существует вид доступа, предназначенный для управления базой данных (администраторы).
3. Интегрированный Web-сервер Domino в зависимости от уровня доступа пользователя отображает доступные ему документы.
4. После выбора пользователем конкретного документа из списка документов сервер запускает конвертер LaTeX2HTML, который преобразует файл в формате TeX в соответствующий набор файлов в формате HTML, доступный для браузеров, который отправляется клиенту.

При помощи клиента Domino Designer, созданного для разработки приложений Domino, для хранения электронных документов создана база данных на сервере. В силу того, что основой структуры базы данных Domino является логический объект «документ», файл в формате TeX (электронный документ библиотеки) фактически является частью логического «документа» Domino (помимо самого файла в формате TeX также есть возможность доба-

вить комментарий к нему или аннотацию, которые будут являться частью логического «документа» Domino). Авторизация пользователей базы данных реализована при помощи встроенных средств администрирования Domino, которые позволяют администратору базы данных создавать и контролировать группы пользователей, а также дают возможность самим пользователям регистрироваться в системе. Само администрирование базы данных при этом осуществляется при помощи клиента Domino Administrator и через Web-браузер.

Кроме того, при помощи программы Designer разработаны несколько форм приложения с информационными полями, представляющие собой графический Web-интерфейс для работы с базой данных.

Все созданные формы, представления, обработчики событий и информационные поля связаны между собой при помощи интегрированного Web-сервера Domino, который также отвечает за связь самого приложения с базой данных документов. Вызов конвертера LaTeX2HTML, преобразующего документы в формате TeX в HTML-страницы, происходит в одном из обработчиков событий базы данных (просмотре документа). При разграничении прав доступа различных категорий пользователей используется встроенный механизм авторизации пользователей Lotus Domino.

В заключение отметим, что оба описанных программных продукта представляют собой выпускные работы авторов, являясь, в то же время, законченными программными продуктами, и с нашей точки зрения, могут быть использованы для создания некоммерческих научно-технических электронных библиотек, содержащих документацию в формате LaTeX.

## Литература

- [1] Гулич С., Гундавара Ш., Бирзекс Г.. CGI программирование на Perl, М.: Символ-Плюс, 2001. – 468 с.
- [2] Гуссен М., Миттельбах Ф., Самарин А.. Путеводитель по пакету LaTeX и его расширению LaTeX2e, М.: Мир, 1999. – 606 с.
- [3] Кирх О. Linux для профессионалов. Руководство администратора сети: Пер. с англ. - 2-е изд. - СПб.: Питер, 2001. - 368 с.
- [4] Линд Д., Керн С. Lotus Notes и Domino5/6. Энциклопедия программиста. 2-е издание: Пер. с англ. М. : ООО «ТИД «ДС», 2005. 1024 с.

## Two Ways of Developing Scientific and Technical Digital Library in TeX Format

© M.O.Kovalenko, A.V.Tsurikov, M.K.Chernyshov

In this article we consider two ways of developing scientific and technical digital library in TeX format. A short description of the library data structure and some arguments in favor of chosen development tools are provided.