

**АВТОРСКОЕ ПРАВО В ДЕЯТЕЛЬНОСТИ
ЭЛЕКТРОННЫХ БИБЛИОТЕК**

Сборник статей под редакцией А. М. Цапенко

МОСКВА
ОАО ИНИЦ «ПАТЕНТ»
2012

УДК 347.78:02(082)
ББК 67.404.31я43
А22

Составители:
Е. А. Данилина, А. М. Цапенко

А22 Авторское право в деятельности электронных библиотек : сборник статей / [сост. : Е. А. Данилина, А. М. Цапенко] ; под ред. А. М. Цапенко. – М. : ИНИЦ «ПАТЕНТ», 2012. – 120 с. : ил. ISBN 978-5-91808-082-5.

И. Данилина, Елена Александровна, сост.

В сборнике анализируются различные авторско-правовые проблемы, возникающие при создании, функционировании и использовании электронных библиотек. Последовательно рассматриваются проблематика библиотечной деятельности на современном этапе развития общества; современное авторско-правовое и библиотечное законодательство; особенности правового режима научных и учебных произведений; конкретные вопросы защиты прав авторов, в том числе вопросы, связанные с так называемым интеллектуальным пиратством, «орфанными» произведениями в фондах библиотек, ограничениями авторско-правовой сферы в гуманитарных целях; проблемы свободных лицензий (лицензии Creative Commons); юридические и морально-этические проблемы функционирования технических средств защиты авторских прав (средств DRM); возможности государственной регистрации и патентования электронных библиотек. Авторы статей – ведущие специалисты в области авторского права, информационных и библиотечных технологий.

Книга предназначена для специалистов в области библиотечного дела, юристов, создателей электронных библиотек и баз данных.

УДК 347.78:02(082)
ББК 67.404.31я43

ISBN 978-5-91808-082-5

© Коллектив авторов, 2012
© Данилина Е. А., Цапенко А. М., составление, 2012
© Оформление. ОАО ИНИЦ «ПАТЕНТ», 2012

1. ЭЛЕКТРОННЫЕ БИБЛИОТЕКИ В СОВРЕМЕННОМ МИРЕ

*Т. С. Маркарова, кандидат философских наук, доцент,
директор ФГНУ «НПБ им. К. Д. Ушинского» РАО*

*А. М. Цапенко, заместитель директора по научной работе
ФГНУ «НПБ им. К. Д. Ушинского» РАО*

Состав и способы представления информационных объектов в современной электронной библиотеке

В современной электронной библиотеке информационные объекты могут быть представлены несколькими способами. Это обусловлено тем, что электронные библиотеки должны обладать рядом принципиально важных достоинств, среди которых широкие возможности по работе с информацией и, в первую очередь, возможность проводить информационные поиски. Последнее особенно ценится в науке, образовании и других сферах деятельности, где требуется работа с большими объемами информации. При этом большое значение имеет качество представления информации: для любого ученого, исследователя или юриста принципиально важным условием является отсутствие искажений в предоставляемой пользователю публикации.

Вместе с тем понятно, что такой объект как книга представляет собой достаточно сложное образование, несущее в себе больше информации, чем содержится в ее основном тексте. Для специалистов в ряде случаев состав, структура издания и некоторые сопроводительные тексты имеют даже большее значение, чем основной текст.

В отличие от традиционных библиотек в электронных библиотеках читатель обращается не к печатному изданию, а к его электронному образу, который получается в результате обработки этого издания. Достаточно часто в реально работающих электронных библиотеках электронные образы не в полной мере соответствуют воспроизводимому оригиналу: отсутствуют фрагменты текста, формулы и таблицы искажены, в текст внесены грамматические ошибки. Все это снижает доверие пользователя к предоставляемой ему информации и к самой электронной библиотеке. В итоге сокращается количество читателей, и электронные библиотеки становятся менее востребованными. Таким образом, затраты на их создание и ведение, являющиеся, как правило, весьма значительными, оказываются неоправданными. Поэтому одной

из главных задач, стоящих перед разработчиками и создателями электронной библиотеки, является точность воспроизведения в ней исходной информации.

На сегодня существуют три основных способа представления информации в электронных библиотеках: графический, символьный и комбинированный. В первом случае печатное издание представляется в виде совокупности графических образов страниц, что обеспечивает высокую точность воспроизведения оригинала. При этом отсутствуют многие возможности по работе с текстами; в частности, нельзя проводить лексический поиск. Графическое представление информации реализуется на основе форматов pdf, djvu, jpg и др.

При символьном способе напечатанный текст воспроизводится как последовательность алфавитно-цифровых знаков той или иной кодовой системы, например Unicode. Это позволяет оперировать с произвольными, заранее не определенными фрагментами представленного текста, что в свою очередь позволяет реализовать разнообразные функциональные возможности. Однако при символьном представлении возникают серьезные проблемы, во-первых, с визуализацией отдельных фрагментов, например математических или химических формул, а во-вторых, с точным воспроизведением исходного оригинала. Для решения указанных проблем используют языки разметки, таблицы стилей, системы преобразований, применение которых заметно повышает сложность подготовки информации. К сожалению, все это не позволяет преодолевать некоторые трудности. При символьном способе представления информации используются такие форматы как html, TEX, fb2.

Стремление к устранению противоречия между точностью воспроизведения информации и реализацией функциональных возможностей привело к созданию комбинированного способа представления информации в электронном виде. Суть данного способа заключается в том, что в информационной системе (электронной библиотеке) информация представляется как в графическом, так и в символьном виде. При этом возможны два варианта. В первом варианте оба вида представления информации доставляются одним файлом (например, pdf_T, djvu_T¹). Во втором (многофайловом) варианте оба вида представления информации доставляются двумя разными способами: графический образ представлен в виде комплекта изображений страниц в формате jpg, а текстовый образ – в виде единого текстового файла, возможно, в формате html. Между файлами, входящими в многофайловое представление, устанавливается определенная связь, которая существенна для навигации и отображения в графическом образе фрагментов текста, найденных в результате поиска. Достоинства ком-

¹ Нижний индекс «Т» у обозначения формата подчеркивает, что речь идет о документах, имеющих текстовый слой.

бинированного способа представления информации очевидны: с одной стороны, пользователь получает точный (графический) образ оригинального издания, с другой – имеются основания для реализации важных функциональных возможностей на базе символьного представления.

Во втором и третьем способах представления информации наиболее трудоемкой является операция вычитки распознанного текста, поскольку она требует ручного труда квалифицированных специалистов и занимает достаточно много времени по любым действующим нормам. В случае символьного представления вычитка играет принципиальную роль и требует особой тщательности, так как именно этот текст представляет оригинал. В случае комбинированного представления требования к вычитке могут быть заметно ослаблены – читатель получает точный графический образ оригинала; влияние же отдельных ошибок на результаты лексического поиска не столь велико, как может показаться.

Таким образом, можно сделать вывод, что наиболее дешевым способом представления информации является графический способ, а наиболее дорогим – символичный. Комбинированный способ занимает промежуточное положение.

В настоящее время именно рассматриваемый выше подход в крупном масштабе используется компанией Google при формировании своей электронной библиотеки. В данном проекте компания сотрудничает с крупнейшими библиотеками США и Европы. С момента создания электронной библиотеки Google в 2005 г. по этой технологии обработано свыше 10 млн книг (режим доступа: <http://books.google.com/>).

Таким образом, можно сделать вывод, что в современной электронной библиотеке информация должна быть представлена следующим образом. Оригинальное издание должно быть представлено в графическом виде, например в формате jpg. Его текст должен быть представлен в символическом виде, например в формате html с минимальной разметкой и упрощенной вычиткой. Между графическим и символическими образами текста устанавливается однозначная связь, которая формируется в процессе создания электронной книги, например посредством специального индексного файла.

Информационные объекты электронной библиотеки можно классифицировать по виду информации и по форме исходного оригинала.

В электронной библиотеке могут быть представлены электронные образы книг или иных текстовых документов, а также изобразительные, аудио- и видеоматериалы.

Текстовая информация является главной составляющей информационного фонда электронной библиотеки, по крайней мере, в большинстве случаев. Основными источниками текстовой информации являются книги, брошюры, журналы, газеты, диссертации и авторефе-

раты диссертаций, нормативные документы, классификаторы, архивные документы, описи архивных фондов, биографические статьи, справки и некоторые другие материалы. Все указанные объекты должны быть объектами атрибутного поиска. Подавляющее их большинство должно участвовать в лексическом поиске, исключением являются отдельные публикации и архивные документы, не допускающие распознавания и представляемые факсимильно в виде изображений, например рукописные тексты.

Изобразительная информация, представляемая в электронной библиотеке, подразделяется на три категории. К первой категории относятся изображения, являющиеся неотъемлемой составляющей публикуемых изданий, ко второй категории – иллюстративный материал, содержащийся в издании и подготовленный, как правило, нетитульным автором этого издания. Третья категория – изображения, являющиеся самостоятельными документами: графические и живописные произведения, фотографии, относящиеся к тематике электронной библиотеки.

Аудио- и видеоинформация может быть представлена фонограммами, фильмами, в том числе анимационными, и видеозаписями, которые являются самостоятельными произведениями. Это могут быть, например, записи лекций, выступлений, экспериментов и учебные материалы.

*А. Б. Антопольский, доктор технических наук, профессор,
главный научный сотрудник Межведомственного
суперкомпьютерного центра РАН*

Опыт и проблемные вопросы создания электронных библиотек

В начале 80-х гг. XX в. произошел прорыв в развитии библиотечного дела – были осуществлены первые попытки создания электронных библиотек.

В 1992 г. конференция Национального научного фонда США (National Science Foundation, NSF) положила начало использованию понятия «цифровая библиотека». В 1990-х гг. в США и странах Европы начали активно разрабатывать программы развития электронных библиотек, а в 1995 г. Европейская комиссия выдвинула создание электронной библиотеки в число приоритетов. При этом наглядно проявилась тенденция к развитию национальных программ по созданию электронных библиотек, в том числе и в форме мультимедийных культурно-образовательных проектов. Среди таких проектов можно отметить следующие: «Память Америки», «Память Испании», «Память мира» (Чешская Республика), SCRAN (шотландская сеть ресурсов по культурному наследию для поддержки образования, аккумулирующая ресурсы музеев, архивов и библиотек). Отличительная черта