

РЕТРОКОНВЕРСИЯ И РЕКАТАЛОГИЗАЦИЯ КАК СРЕДСТВО ОТРАЖЕНИЯ РЕТРОСПЕКТИВНОГО ФОНДА НАЦИОНАЛЬНОЙ БИБЛИОТЕКИ БЕЛАРУСИ

*Зеленевская Наталья Петровна, зав. отделом библиографической
обработки документов НББ*

Главной целью библиотеки, как известно, является обеспечение доступа пользователя к необходимым ему документам и данным, независимо от того, в каком виде они представлены и в каком месте они хранятся. Основным направлением развития информационно-поискового аппарата НББ является ритмичное пополнение электронного каталога (ЭК) за счет каталогизации входного документного потока и ретроспективного отражения ранее полученных документов – электронной обработки карточных каталогов и частичной рекаталогизации отдельных фондов.

НББ осуществляет рекаталогизацию национальных документов, периодических и продолжающиеся изданий (в том числе газет), редкой и старопечатной книги. В процессе рекаталогизации библиографические записи создаются на основе просмотра документа *de visu*, индексирование осуществляется по типовой технологии смысловой обработки, что позволяет создать полноценную структурированную запись, практически ничем не отличающуюся от записи, полученной при каталогизации текущего документного потока.

К сожалению, эти работы ведутся недостаточно интенсивно в силу сложности библиографической обработки вышеуказанных документов. В частности, обработка национального документа близится к завершению – составлены 127 146 библиографических записей на весь книжный фонд. В течение 2006–2008 гг. будет завершён процесс смысловой обработки примерно 15 000 документов, после чего рекаталогизация национального фонда будет завершена.

Достаточно сложен процесс рекаталогизации периодических и продолжающихся изданий. Объем этого фонда составляет 2 775 906 экземпляров (из них на иностранных языках – 971 110). Библиографические записи создаются на каждое наименование издания с установленными по “Летописям периодических и продолжающихся изданий” ссылками и связями. Осуществляется не только доработка библиографической информации, но и упорядочивание фондов. В основном, это касается продолжающихся изданий, которые находятся частично в книжном фонде, частично – в фонде периодики, на что есть объективная причина – вид конкретного издания определяла Всесоюзная книжная палата, присылая инвентарные листы с экспедиционными номерами и в дальнейшем отражая библиографические записи в соответствующей “Летописи...”. На сегодняшний день библиографы определяют, в каком фонде должно храниться конкретное издание, и отдел комплектования перенаправляет документы в соответствующие фонды. Так, из книжного фонда в фонд

периодики на протяжении 2003–2006 гг. были переведены 3 642 экземпляра. В результате такой кропотливой работы выявляются излишняя экземплятность либо пробелы в комплектности, дефектность отдельных изданий. В таких случаях специалистами отделов, занимающихся комплектованием и хранением фондов, принимаются все необходимые меры для исправления неточностей, упорядочивания и докомплектования фонда продолжающихся изданий.

Поскольку задачей любой библиографической записи является продвижение к пользователю конкретного документа, необходимо отметить, что все периодические издания подвергаются смысловой обработке, в результате которой формируется поисковый образ документа в виде индексов УДК, рубрик ГРНТИ и ключевых слов. На данном этапе в базе данных локального ЭК создано 22 049 библиографических описаний периодических и продолжающихся изданий на русском, белорусском и украинском языках и 5 390 – на иных языках мира.

Необходимо сказать, что в условиях старого здания библиотеки в ЭК не отражался фонд газет. В течение короткого времени (апрель–сентябрь 2006 г.) в новой АБИС был создан электронный справочник наименований всех имеющихся в фонде библиотеки газет общим количеством 4454. В записях представлена полная библиографическая информация и, дополнительно, указан вид газет – республиканские, областные, районные, эмигрантские и т. д. В дальнейшем библиографические записи будут дополняться сведениями о наличии комплектов газет с идентификацией каждой подшивки и уточнением места ее хранения, что позволит осуществлять заказ этого вида документов через ЭК.

Работы по рекаталогизации фондов редкой и старопечатной книги, рукописей и архивных материалов ведутся планомерно по коллекциям: белорусские документы XVI–XVIII вв., напечатанные латинским шрифтом, миниатюрные издания, кириллица, коллекция инкунабулов, русская эмигрантская печать и т. д. Заканчивается работа с коллекциями палеотипов, эльзевиров, книг русской гражданской печати и петровского времени.

Тем не менее, существенную часть фонда библиотеки (около 4 000 000 названий) подвергать рекаталогизации не представляется возможным по причине больших трудо- и временных затрат. Перед библиотекой была поставлена задача конверсии Генерального алфавитного каталога. Объектом ретроконверсии был определен массив карточек (примерно 3 800 000 единиц), отражающий по языковому и видовому признакам фонд, каталогизированный на карточках до начала внедрения АБИС. Поскольку в алфавитном каталоге не отражен фонд картографических изданий, в работу был включен и Систематический каталог картографических изданий. Сканирование каталогов было начато с русскоязычной части Генерального служебного алфавитного каталога книг (ГСАК) в марте 2004 г. и закончено в ноябре 2006 г.

Результатом проведения сканирования явился “Имидж-каталог”, в котором размещены 3 988 315 графических образов. В предлагаемом пользователям

библиотеки каталоге – массив факсимильных образов лицевых сторон каталожных карточек, отражающих библиографическую информацию, зафиксированную в составных частях ГСАК и Систематическом каталоге картографических изданий. Несмотря на то, что продолжается рекаталогизация национального фонда, с целью полноты представления информации в “Имидж-каталоге” дополнительно была отсканирована белорусскоязычная часть алфавитного каталога. Таким образом, имидж-каталог дает представление о наличии в фонде библиотеки книг, брошюр, авторефератов диссертаций и диссертаций, альбомов, атласов, электронных и картографических ресурсов, нот и изодокументов, изданных с 1800 г. по 2005 г., и данные о местонахождении всех экземпляров документа в библиотечном фонде (его шифры хранения). Структура имидж-каталога повторяет структуру соответствующих карточных каталогов: этикетки и нумерационное обозначение на каталожных ящиках, разделители внутри ящика – буквенные, слоговые, авторские, классификационные индексы, ссылочные, отсылочные и добавочные карточки. Просмотр карточек каталога производится путем перелистывания образов.

Графические образы каталожных карточек являются источником представления структурированной библиографической записи. Работа с графическим массивом производится по следующей технологии:

1. распознавание образов карточек и преобразование текста с графического источника;

2. загрузка в базу данных информационно-технологической системы ретроконверсии (ИТСР) графической информации и, затем, текстовой информации, полученной после распознавания;

3. корректировка текста методом сравнения с информацией графического образа каталожной карточки. Для обеспечения минимизации ошибок в тексте операция выполняется двумя разными операторами с последующим автоматическим посимвольным сравнением;

4. маркировка текста, т.е. выделение в тексте карточки полей библиографической записи согласно формату BelMARC;

5. выгрузка информации из базы данных ИТСР в выходные файлы, структурированные в соответствии с ГОСТ 7.14–98 (ISO 2709–96).

Таким образом, обработка массивов построена на комбинировании ручных и автоматических методов. Для каждого типа полей производится группировка работ по операторам (например, заполнение числовых полей – шифров и индексов, маркировка полей, редактирование записи), что позволяет более эффективно и качественно обрабатывать документы, а вся вводимая информация формализуется правилами ввода. С целью проверки правильности представления и обеспечения полноты имеющихся на карточке данных при работе с текстом применяется метод верификации, т.е. только при условии полного совпадения обработанной двумя операторами информации, далее она передается на следующую технологическую операцию, а при наличии рассогласования администратор должен принять решение – возврат текста на повторную обработку оператору либо самостоятельное устранение рассогласования.

Итак, в целом, по каждой каталожной карточке выдержана следующая технология:

1. автоматизированное сканирование записи каталожной карточки и формирование ее графического образа;

2. специальная программная обработка графического образа каталожной карточки с автоматическим распознаванием полей данных и формирование электронного текста карточки;

3. анализ оператором электронного текста карточки, выявление полностью нераспознанных карточек или отдельных нераспознанных элементов с последующим их восстановлением с клавиатуры компьютера;

4. подключение опытных библиографов для принятия решения в спорных ситуациях.

Описанная технология, предусматривающая режим верификации при выполнении самых ответственных операций, позволяет выполнять довольно жесткие требования к качеству создаваемых библиографических записей. Заложенные в ИТСП технологические принципы обработки и контроля гарантируют возможность выгрузки в выходной файл только полностью завершенных блоков данных, кратных объему каждого из каталожных ящиков. Одновременно с выходным файлом предоставляется дополнительная информация, позволяющая контролировать полноту обработки массивов.

Несмотря на финансовые и трудовые затраты, длительность и сложность проведения ретроспективной конверсии и рекаталогизации в результате библиотекой решаются такие глобальные проблемы, как:

– обеспечение сохранности библиографической информации, зафиксированной и хранящейся на карточках в каталоге;

– уточнение библиографических данных и создание библиографического описания по правилам действующего ГОСТа для всех карточек;

– доступность к справочно-библиографическому аппарату неограниченного количества пользователей (как в самой библиотеке, так и возможность удаленного доступа через сеть Интернет, причем круглосуточного);

– получение электронных копий каталожных карточек;

– возможность удаленного заказа документов.

В результате проведения рекаталогизации и ретроконверсии карточных каталогов создается исходная основа будущего полноценного электронного справочно-поискового аппарата на весь фонд библиотеки. Обработанный должным образом ретроспективный библиографический массив в будущем станет базой для корпоративного взаимодействия библиотек по созданию и использованию электронных информационных ресурсов страны.