

**Лавренова О.А.**

*Российская государственная  
библиотека, Москва, Россия*

## **Создание электронных каталогов в Российской государственной библиотеке: стратегия, лингвистика, проблемы**

1. Автоматизация в РГБ ведется в двух направлениях: в библиотеке работает АСНТИ "Информкультура" и по сути выполняются функции информационного органа по культуре и искусству в РФ, а также создается автоматизированная информационно-библиотечная система (АИБС) РГБ, состоящая из различных банков данных, связанных между собой единой концепцией автоматизации в плане технологии, информационного и лингвистического обеспечения.

Разумеется, быстрое развитие электронной вычислительной техники, машинных носителей и вместе с ними методов обработки и передачи информации требует периодического пересмотра концептуальных положений, смены акцентов в организации работ и в особенности переработки структуры комплекса технического обеспечения системы, которые принципиально меняются в мировой практике через каждые 2—3 года.

Электронные каталоги библиотеки являются ведущими банками данных в рамках системы.

В РГБ предполагается создать следующие электронные каталоги (ЭК):

а) ЭК книг, нот, карт, авторефератов и продолжающихся изданий на русском языке, языках народов РФ и иностранных языках, использующих кириллический и латинский алфавиты и имеющих в фонде РГБ; вначале это должен быть ЭК текущих поступлений с определенной даты, а затем в него вливаются ретроспективные каталоги;

б) электронные каталоги на языках, использующих другую графику;

- в) сводный ЭК русской книги 1801—1825 гг.;
- г) сводный ЭК русской книги 1826—1917 гг.;
- д) каталоги архивных фондов РГБ (например каталог архивов Чехова, Достоевского);
- е) сводные ЭК иностранных нотных и картографических изданий; они могут существовать и в рамках ЭК (а);
- ж) ЭК защищенных диссертаций;
- з) ЭК баз данных и программных продуктов, имеющих в РГБ;
- и) ЭК славяно-русских рукописных книг и произведений из них;
- к) электронные каталоги рукописных книг на других языках;
- л) электронные каталоги периодических изданий;
- м) электронные каталоги статей и произведений из газет и журналов.

Ретроспективная конверсия традиционных каталогов РГБ в машиночитаемую форму требует огромных затрат, но решение этой задачи нельзя откладывать, поскольку необходимо обеспечить сохранность информации, содержащейся в карточных каталогах, и доступность ее для читателей различных библиотек РФ и других государств.

В первую очередь следует перевести на машинные носители следующие алфавитные каталоги:

- сводные ЭК русской книги XIX века;
- генеральный алфавитный каталог книг на русском языке в РГБ, исключая библиографические записи (БЗ), относящиеся к первым двум каталогам;
- генеральный алфавитный каталог периодических изданий на русском языке;
- каталог русской книги до XVIII века (имеется книжный вариант).

Затем имеет смысл перевести в машиночитаемую форму: ретроспективный сводный каталог зарубежных журналов по общественным и гуманитарным наукам за 250 лет (существует в машинописном виде);

- каталоги редких иностранных книг;
- каталоги отечественных и иностранных карт и нот.

Перевод в машиночитаемую форму генерального алфавитного каталога на иностранных европейских языках трудно признать экономически оправданным без использования зарубежных машиночитаемых библиографических баз данных.

Перевод в машиночитаемую форму карточных каталогов может быть осуществлен только путем набора с клавиатуры компьютера, но не путем сканирования, так как текст БЗ на карточках невысокого качества, часто рукописный. Перевод на машинные носители отдельных каталогов, существующих в книжной форме, напротив, возможен по программам сканирования и распознавания образов.

В качестве новой задачи следует рассматривать создание таких банков данных по коллекциям РГБ, в которых сочетаются БЗ и факсимильные изображения страниц. Такие базы данных формируются путем клавиатурного ввода БЗ и сканирования страниц с изображениями изданий, газет, фотографий и т.д. Первый опыт РГБ — банк данных "Автографы известных деятелей России". Имеется также зарубежная разработка банка данных "Плакаты".

В настоящее время новая иностранная литература каталогизируется на старой отечественной мини-ЭВМ СМ. Законченные записи выводятся на магнитные ленты для изготовления карточек на ЭВМ ИР 3000 и лазерном принтере для разных каталогов, часть записей передается на МЛ в сводный каталог иностранных книг по естественным наукам, технике, медицине, сельскому хозяйству в ГПНТБ России. Поддерживается база данных на ЭВМ ИР 3000/70 с элементами поиска.

Задача обработки текущих поступлений отечественных изданий автоматизирована частично: ведется ввод БЗ только по авторефератам диссертаций и отечественным нотам и картам для подготовки изданий.

Ретроспективная конверсия в электронную форму традиционных каталогов реализуется на персональных компьютерах для двух сводных каталогов "Русская редкая книга 1801—1825 гг." и "Русская книга 1826—1917 гг.", формируются электронные каталоги, но очень медленно, так как не хватает средств для оплаты ввода. Создается ЭК славяно-русских рукописных книг из фонда РГБ и произведений из них.

Следует упомянуть также подготовку проблемно-ориентированных баз данных по различным темам.

На следующем этапе планируется создание многотерминальных вычислительных комплексов на базе персональных компьютеров, формируется технология загрузки и доработки массивов БЗ отечественного обязательного экземпляра с дискет РКП в ЭК на персональном компьютере, а также технологии довода в этот ЭК БЗ покупаемых РГБ изданий и предоставления этого ЭК в пользование читателям.

Затем необходим перевод обработки поступлений всех изданий на современную высокопроизводительную вычислительную технику и формирование на ней системы электронных каталогов различных видов изданий. На персональных компьютерах остаются отдельные локальные исследовательские ЭК, в которых содержатся БЗ особой структуры (например БЗ редкой книги или рукописей).

Для поддержки электронных каталогов (а), (е), (ж), (л), (м) с подсистемой обслуживания приобретено программное обеспечение VTLS (США), которое реализует все основные функции ЭК с подсистемой обслуживания читателей. Совместно с фирмой VTLS, Inc. проведена адаптация системы к задачам РГБ и сформирована ее русская версия. В настоящее время фирма занимается устранением имеющихся недостатков в этой версии системы, доработкой недостающих функций. Препятствием для внедрения служат финансовые сложности. Рассчитываем на финансирование автоматизации из "нетрадиционных" источников.

2. Лингвистика электронных каталогов — это в первую очередь формат представления элементов библиографических записей и информационно-поисковый язык (ИПЯ) для тематического поиска. По сути лингвистика ЭК и есть та основа, которая обеспечивает единство различных каталогов или, напротив, их разъединяет. Поэтому сотрудничество различных библиотек в плане компьютеризации имеет смысл осуществлять преимущественно в данной области.

В РГБ все ЭК должны быть основаны на использовании единого формата представления элементов БЗ — русской версии USMARC-формата (RUSMARC). В ЭК идентифицируются все элементы библиографических записей для тех видов документов,

которые хранятся в РГБ, а также данные, необходимые для подготовки библиографических изданий с указателями и каталожных карточек в соответствии с действующими в РФ стандартами. Ведется специальный банк данных "Формат" для поддержки русифицированного формата MARC.

Для тематического поиска в большинстве ЭК РГБ должны использоваться таблицы библиотечно-библиографической классификации (ББК), разработанные и совершенствуемые РГБ.

Сложилась странная ситуация, когда библиотеки РФ требуют машиночитаемые таблицы ББК, чтобы в своих электронных каталогах обеспечить принятые в библиотеках принципы тематического поиска с новыми возможностями, но в то же время сталкиваешься со скептическим отношением к ББК со стороны людей, поверхностно знакомых с этой классификацией, но принимающих решения при разработке путей компьютеризации тех или иных библиотек. В какой-то мере это можно понять. Достаточно прочесть несколько страниц введения к полному изданию таблиц — и окунаешься в ту идеологическую атмосферу, от которой только что отвык. Однако стоит преодолеть эту слабость и вникнуть в структуру и содержание таблиц. Беспристрастный специалист в области информационного поиска непременно почувствует, что эти стандартные, идеологически выверенные фразы введения и соответствующие разделы в первых рядах были стандартным прикрытием нормального, честного труда хороших мастеров классификации.

Каждый кандидат или доктор наук, защитившийся не в последние годы, аналогичным образом выстраивал цитаты и ссылки в своих диссертациях, от чего научная ценность основного содержания ничего не потеряла и не прибавилась. В ББК пострадали отдельные разделы: "Экономика", "История", "Государство и право" и т.п. В настоящее время все они перерабатываются в рамках принципиальной структуры таблиц. В то же время разделы техники и технических наук, медицины, разделы для нотных и картографических изданий и т.д. не имели идеологической интерпретации. Они перерабатываются или редактируются только с точки зрения современного развития наук.

Разработчики лингвистики ЭК ГЛ (РГБ) не принадлежали к приверженцам ББК или УДК, а создавали прежде всего деск-

рипторные языки с тезаурусами. Именно поэтому мы не обольщались планами создания тезаурусов для универсальной библиотеки, зная трудоемкость таких работ.

Выбор ББК в качестве ИПЯ ЭК был обусловлен следующими факторами:

ББК используется в традиционных каталогах большинства библиотек страны;

ББК признается в мировой библиотечной теории и практике как одна из равноправных универсальных классификаций;

УДК как альтернативный вариант не может быть использована, поскольку немыслимо в одночасье и даже за несколько лет перевести ни РГБ, ни другие солидные библиотеки на новые методы классификации; кроме того, таблицы УДК структурированы, типизированы менее логично и строго, чем ББК, и кое-что черпают сейчас из нашей классификации;

ББК при определенном редактировании легко преобразуется в ИПЯ с иерархическими и ассоциативными связями и грамматикой;

в отличие от других классификаций существует несколько версий ББК: полные таблицы для научных библиотек, варианты для областных и массовых библиотек и т.д.;

таблицы ББК постоянно развиваются, причем отечественными авторами, в то время как распространение УДК и изменения в этих таблицах требуют выполнения международных условий и определенных затрат.

Создание в РГБ машиночитаемых таблиц ББК с предметным входом (доступом) резко подняло реноме классификации в глазах различных библиотек, которые живо заинтересованы в приобретении файлов ББК на дискетах для загрузки в свои электронные каталоги. Никому не хочется повторять эту тяжелую работу.

Деления ББК в машиночитаемых таблицах представляются в форме нормативных записей (Authority Records). Обеспечиваются поиск по наименованиям делений, индексам, предметному входу (Subject access), учет иерархических и ассоциативных отношений между делениями, вывод ссылок и методических указаний (Information Notes). Приблизительно 1/4 таблиц уже сформирована на машинных носителях, причем одновременно создаются ма-

шиночитаемые таблицы ББК для областных библиотек. Ряд библиотек уже приобрели файлы из полных и областных таблиц.

Кроме ББК, в виде Authority files (файлов нормативных записей) должны быть представлены на машинных носителях записи об индивидуальных и коллективных авторах (совместно с РКП). Для проведения этих работ в РГБ создается Служба ведения нормативных записей, ведется подготовка русской версии формата USMARC/Authorities. Подобрана и обрабатывается документация по организации подобных работ за рубежом. Совместно с РКП и ГПНТБ России переведен на русский язык формат UNIMARC/Authorities.

3. Проблемы создания ЭК в РГБ связаны, как обычно, в основном с финансированием.

При кажущемся многообразии вычислительной техники в библиотеке у нее никогда не было средств на приобретение комплекта, необходимого для создания главной части АИБС — электронного каталога библиотеки с подсистемой обслуживания. Может быть внедрена заведомо упрощенная система с ограниченными возможностями. Имеется для ЭК только одна ЭВМ HP 3000/70, недостаточно дисплеев, принтеров, во многих помещениях невозможно проложить каналы связи и установить вычислительную технику. Один компьютер с памятью 2 Гб — очень мало для ЭК национальной библиотеки. Отключение этой машины, которая уже плохо работает, будет означать прекращение обслуживания, так как нет замены. Необходима современная техника и дублирующие вычислительные ресурсы для поддержки системы в постоянном рабочем режиме.

В заключение перечислим задачи, решение и финансирование которых необходимо в интересах не только РГБ, но и большинства других библиотек РФ.

Разработка проекта создания системы сводных каталогов фондов библиотек России (в первую очередь электронных) и ее реализация, согласование проблем с Миннауки и РАН; обеспечение централизованной каталогизации в РФ.

Инициативная группа специалистов РГБ разработала проект концепции системы сводных каталогов фондов библиотек России. В этой концепции изложены основные принципы построения системы, предложена определенная интерпретация понятий, связан-

ных с данной проблемой. Документ предложен для обсуждения заинтересованными специалистами, чтобы затем приступить к разработке проекта системы, которая должна включать сводные каталоги по различным видам документов на разных языках и по всем периодам времени.

Перевод в машиночитаемую форму генеральных алфавитных каталогов РГБ и перевод их на оптические диски для всех библиотек.

Формирование технологии CD-ROM в России.

Упорядочение приобретения за счет бюджетных средств библиографических баз данных на CD-ROM за рубежом в рамках иностранного комплектования библиотек.

Программа "Память России" (только начато финансирование).

Завершение работ по формированию машиночитаемых таблиц ББК.

Создание службы ведения нормативных записей.

Стандартизация коммуникативного формата РФ (на базе формата USMARC) и русской версии форматов MARC для библиографических и нормативных записей.

Создание сетей передач данных между библиотеками, выход в международные сети.