

ИНФАРМАЦЫЙНЫЯ РЭСУРСЫ І ИНФАРМАЦЫЙНЫЯ ТЭХНАЛОГІ У БІБЛІЯТЭКАХ

Л.Л.Астапович, *заведуючы аддзелам
научнай сістэматызацыі дакументаў;
Т.С.Ерэма*, *заведуючы сектаром
класіфікацыйнага індэксавання дакументаў,
Цэнтральная навучная бібліятэка
ім. Я.Коласа НАН Беларусі*

СМЫСЛОВАЯ ОБРАБОТКА ДОКУМЕНТОВ ПРИ ВЕДЕНИИ ЭЛЕКТРОННОГО КАТАЛОГА В ЦЕНТРАЛЬНОЙ НАУЧНОЙ БИБЛИОТЕКЕ им. Я.КОЛАСА НАН БЕЛАРУСИ

Смысловая обработка документов - это один из наиболее сложных процессов каталогизации. В условиях компьютеризации библиотечного дела при ведении электронного каталога (ЭК) возрастает ответственность индексаторов за этот процесс. В Центральной научной библиотеке им. Я.Коласа Национальной академии наук Беларуси (ЦНБ НАН Беларуси) при смысловой обработке документов для выражения их содержания используются информационно-поисковые языки (ИПЯ) классификационного типа (универсальная десятичная классификация - УДК) и вербального (ключевые слова - КС - и дескрипторы). Применение УДК как ИПЯ расширяет возможности поиска в ЭК и обеспечивает лингви-

стическую совместимость ЭК ЦНБ НАН Беларуси с ЭК библиотек, использующих УДК.

Ранее, когда основным источником поиска литературы по теме являлся систематический каталог (СК), ведущее значение принадлежало УДК. Но с организацией в библиотеках ЭК большое значение стало придаваться координатному индексированию документов с использованием КС и дескрипторов.

На начальном этапе ведения ЭК в ЦНБ НАН Беларуси координатное индексирование при смысловой обработке документов осуществлялось с помощью КС. Проводились разработка методики координатного индексирования документов и работа по составлению перечня КС в виде словаря. В первое время при координатном индексировании документов контроль КС полностью отсутствовал. Индексаторы, не обладавшие опытом определения КС и не имевшие методики по их составлению, выбирали ключевые слова непосредственно из текста документа без учета того, какие КС уже использовались ранее для индексирования таких же или близких по смыслу документов. В данном случае не устранились синонимия, полисемия и омонимия КС, не приводились к избранному виду их грамматические формы. Недостатки такой работы очень быстро проявились при формулировке запросов для поиска в ЭК библиотеки. При выработке единого подхода к составлению КС специалисты ЦНБ НАН Беларуси консультировались с сотрудниками других библиотек. С использованием опыта библиотек по составлению КС, применением рубрик из таблиц УДК и АПУ к СК, терминов из различных словарей и тезаурусов был разработан подход к составлению КС, утверждена "Методика составления КС" для ЭК ЦНБ НАН Беларуси на документы, изданные на русском и белорусском языках (1996) [1].

В 2003 г. ЦНБ НАН Беларуси перешла на систему АБИС БИТ-2000и, предусматривающую интерактивную работу с онлайн-каталогом библиотеки и со сводным электронным каталогом (СЭК) библиотек Беларуси, и на Национальный обменный формат BELMARC, соответствующий Международному коммуникативному формату UNIMARC и Российскому формату RUSMARC [3].

Для отражения тематики документов как в текстовом виде, так и с помощью условных обозначений в формате BELMARC предназначен 6 блок определения тематики. Поля этого блока дают возможность применять широкий спектр языков классификационного и дескрипторного типа с контролируемой и неконтролируемой лексикой. Каждое из полей используется для определенного вида предметной рубрики, являющейся точкой доступа.

В отконвертированной базе данных ЭК библиотеки индексы УДК были занесены в поле 675 (УДК), а ключевые слова - в поле 610 (неунифицированные тематические термины). Составление КС осуществлялось по ранее разработанной методике. Продолжилась работа со словарем КС - тематическим словарем ЭК, проводились контроль состава словаря, нормализация КС, устранение омонимии и полисемии осуществлялось при помощи помет (лексиграфически). В результате образовался словарь ключевых слов, представляющий собой их алфавитный перечень, в котором применяется система ссылок и отсылок. После многочисленных доработок и изменений, проведения полной редакции словаря КС в 2001 г. была утверждена переработанная "Методика составления ключевых слов для электронного каталога" и их ввода в ЭК на документы на русском и белорусском языках.

ЦНБ НАН Беларуси располагает значительным фондом естественнонаучной литературы на иностранных языках, который постоянно пополняется новыми документами. При составлении методики координатного индексирования документов на иностранных языках учитывалось, что задачей ЦНБ НАН Беларуси является обеспечение потребности ученых в научной информации. Для многоаспектного выражения смыслового содержания документов по научной тематике, изданных на английском, французском, немецком и польском языках, КС приводятся на русском языке и языке оригинала. На документы, изданные на других иностранных языках, КС даются на русском языке. Эти решения отражены в "Методических рекомендациях по определению и введению ключевых слов в электронный каталог документов на иностранных языках" [5].

При анализе поискового образа документа (ПОД), создаваемого при смысловой обработке в разных библиотеках, стало очевидным, что практика использования только КС не очень эффективна. КС как ненормированная лексика отражает субъективный взгляд индексаторов. Сформулированные ими КС не всегда совпадают с формулировкой запросов пользователей. Поэтому при тематическом поиске в ЭК происходит потеря данных: из-за различной формулировки КС релевантные документы оказываются вне результатов поиска.

В настоящее время ведется работа по созданию системы корпоративной каталогизации изданий (СКК) и ведению СЭК Беларуси. Информационную основу СКК составляет СЭК, который представляет собой базу данных, содержащую библиографические (БЗ) и авторитетные (АЗ) записи, представленные в BELMARC-формате. Записи должны формироваться всеми участниками СКК на основе принятых соглашений. В технологию корпоративного взаимодействия вовлекаются локальные информационные ресурсы библиотек-участниц, в том числе и ЦНБ НАН Беларуси как участницы СКК.

Библиотекам для унификации БЗ необходимо использовать при смысловой обработке документов контролируемую лексику - АЗ. Структура для поддержки функционирования СКК и СЭК создается на базе Национальной библиотеки Беларуси (НББ), что позволяет при координатном индексировании документов библиотекам-участницам использовать АЗ-дескрипторы, создаваемые НББ для Национальной базы данных АЗ и положенные в основу тематического словаря СЭК [4].

Для формирования ПОД при смысловой обработке в ЦНБ НАН Беларуси АЗ-дескрипторы применяются для следующих наименований предметов:

- 600 Имя лица, используемое как предмет;
- 601 Наименование организации, используемое как предмет;
- 602 Родовое имя, используемое как предмет;
- 604 Имя автора и заглавие документа (произведения), используемые как предмет;

605 Заглавие, используемое как предмет;

606 Наименование темы, используемое как предмет;

607 Географическое наименование, используемое как предмет;

608 Форма (вид), жанр, физические характеристики документа;

616 Наименование торговой марки (товарного знака), используемое как предмет [2].

После прохождения обучающего тренинга в НББ в мае 2009 г. сотрудники библиотеки приступили к работе по формированию АЗ, в том числе и для блока определения тематики. Особое значение имеет возможность создания АЗ на энтитет "Тематический предмет". Это сделает координатное индексирование документов эффективнее, позволит более полно раскрывать содержание документов, особенно по естественнонаучной тематике.

Применяя АЗ из СЭК при смысловой обработке, мы сталкиваемся с разрозненностью использования дескрипторов при индексировании схожих документов в библиотеках-участницах СКК. Поэтому в целях унификации БЗ для СЭК и с учетом того, что в методике координатного индексирования ЦНБ НАН Беларуси и НББ есть значительные различия, вырабатываем новую методику, которая позволит обеспечить точное и полное описание документного потока посредством КС и АЗ-дескрипторов в удобной для пользователей форме.

Постоянно ведется редактирование тематического словаря ЭК библиотеки. Из словаря выводятся КС, заменяемые АЗ-дескрипторами, которые заносятся в соответствующие поля 6 блока определения тематики. Замена КС на АЗ проводится как программно, так и вручную.

Но, помимо АЗ-дескрипторов, при координатном индексировании документов индексаторы будут выбирать и КС. В научной литературе, поступающей в библиотеку, часто содержатся описания новых явлений и понятий, употребляются новые термины. В случае отсутствия этих терминов в отраслевых энциклопедиях, словарях и электронных ресурсах, когда единственным источником информации о предмете является индексируемый документ, допускается использование КС.

Использование 4 версии УДК и сформированных по единым принципам тематических рубрик, особенно АЗ, дает большие возможности для осуществления смысловой обработки документов, создания более точного ПОД, унификации БЗ, а также поможет пользователям ориентироваться в локальном ЭК и в СЭК, проводить в них тематический поиск.

1. *Астапович, Л.Л.* Проблемы и тенденции формирования средств лингвистического обеспечения в ЦНБ им. Я.Коласа Национальной академии наук Беларуси / Л.Л.Астапович, О.М.Дрозд // Корпоративные библиотечные системы: технологии и инновации: тр. VI науч.-практ. конф., 23-29 июня 2008 г., Санкт-Петербург (Россия), Хельсинки, Тампере (Финляндия), Стокгольм, Упсала (Швеция) / Санкт-Петербург. гос. политех, ун-т [и др.]. - СПб., 2008. - С. 127-132.

2. *Белорусский коммуникативный формат представления библиографических записей в машиночитаемой форме (BELMARC) [Электронный ресурс]* / М-во культуры Респ. Беларусь, Нац. б-ка Беларуси. - Мн., 2006. - Режим доступа: http://old.nlb.by/html/news2006/30_08_06/BELMARC.pdf. - Дата доступа: 18.06.2009.

3. *Дрозд, О.М.* Внедрение национального коммуникативного формата BELMARC в работу ЦНБ им. Я.Коласа Национальной академии наук Беларуси / О.М.Дрозд, Л.Л.Астапович, М.Ю.Грядовкина // Научные и технические библиотеки. - 2007. - № 7. - С. 39-42.

4. *Кузьминич, Т.В.* Смысловая обработка документного потока в условиях корпоративного взаимодействия библиотек Беларуси / Т.В.Кузьминич, С.А.Пугачева // Библиотеки и информационные ресурсы в современном мире науки, культуры, образования и бизнеса: материалы 14-й Междунар. конф. "Крым-2007" [Электронный ресурс]. - М., 2007. - 1 электрон. опт. диск (CD-ROM).

5. *Методические рекомендации по определению и введению ключевых слов в электронный каталог документов на иностранных языках каталога* / НАН Беларуси, Центр. науч. б-ка им. Я.Коласа; сост.: Л.С.Солодова, Л.А.Шилова, Я.Е.Шанина; науч. ред. Н.Ю.Березкина. - Мн., 1998. - 14 с.