

3. БИБЛИОТЕЧНО-ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ. ПУБЛИКАЦИОННАЯ АКТИВНОСТЬ

УДК 025.355:[025.43+025.45].02/.03-048.25(476)

РЕДАКТИРОВАНИЕ ЭЛЕКТРОННОГО КАТАЛОГА: ПРОБЛЕМЫ ОБЕСПЕЧЕНИЯ ТЕМАТИЧЕСКОГО ПОИСКА В ЦЕНТРАЛЬНОЙ НАУЧНОЙ БИБЛИОТЕКЕ НАН БЕЛАРУСИ

Л. Л. Астапович

Центральная научная библиотека им. Я. Коласа НАН Беларуси, Минск

Рассмотрены процессы и влияние редактирования электронного каталога на улучшение его поисковых возможностей и качество выполнения информационного поиска. Описана работа по актуализации тематических словарей. Отражены особенности проведения редактирования электронного каталога по классификационным разделам.

Электронный каталог (ЭК) является важным информационным ресурсом библиотеки, раскрывающим состав фонда, обеспечивающим свободный доступ к библиографической информации и реализацию ее многоаспектного поиска. Достаточно востребованным у пользователей является тематический поиск в ЭК с использованием классификационных и координатных информационно-поисковых языков (ИПЯ), которые применяются в библиотеках для смысловой обработки документов.

Смысловая обработка, в ходе которой формируется поисковый образ документов (ПОД), позволяет отражать в ЭК информацию о содержании и назначении документов, образующих фонд библиотеки. Однако для обеспечения корректного поиска и получения результатов, удовлетворяющих тематическим запросам, необходимо проводить редактирование ЭК.

В Центральной научной библиотеке им. Я. Коласа НАН Беларуси (ЦНБ НАН Беларуси) в качестве классификационного ИПЯ принята Универсальная десятичная классификация (УДК), авторитетные записи (АЗ) и ключевые слова (КС) используются как координатные ИПЯ. Соответственно, в библиотеке осуществляется редактирование ЭК по разделам УДК и его тематических словарей.

Координатное индексирование, выполняемое посредством АЗ и КС, позволяет детально и многоаспектно раскрыть тему документа и обозначить ее в ПОД. АЗ создаются библиотеками – участницами системы корпоративной каталогизации (СКК), в том числе и ЦНБ НАН Беларуси, в сводном электронном каталоге (СЭК) библиотек Беларуси. Затем, при смысловой обработке документов, АЗ заимствуются из СЭК и в составе библиографических записей (БЗ) поступают в локальный каталог библиотеки, пополняя тематические словари. Для описания содержания документов и обеспечения полноты поискового образа используются и КС, которые составляются в основном на узкоспециальные и авторские термины, не представленные в авторитетных источниках, и дополняют словарь «Неконтролируемые тематические термины» ЭК.

Участницы СКК ведут постоянную деятельность, направленную на оптимизацию состава словарей СЭК. Учитывая необходимость унификации языковых средств и установления единообразия состава тематических словарей ЭК и СЭК, в ЦНБ НАН Беларуси осуществляется актуализация АЗ для внесения требуемых изменений в словарь каталога библиотеки, контроль этих изменений проводится по вспомогательным информационным таблицам СЭК. Полученные из таблиц сведения используются также

и для корректировки ПОД, которые связаны с преобразованными АЗ. Для улучшения структуры и качества словарей, обеспечения результативного поиска в ЭК в отделе лингвистического обеспечения (ОЛО) ЦНБ НАН Беларуси ведется редактирование следующих словарей: Тематический предмет, Родовое имя, Торговая марка, Географическое название, Унифицированное заглавие, Имя/заглавие, Форма, жанр и физические характеристики документа, Неконтролируемые тематические термины.

Редактирование словарей началось в 2008 г., когда специалисты ОЛО приступили к работе над словарем «Неконтролируемые тематические термины». Большинство КС дублировали АЗ или являлись их синонимами. Это в значительной степени негативно влияло на поиск, не позволяло выявить все библиографические ресурсы по интересующей пользователя тематике. Необходимо было КС заменить уже созданными АЗ или сформировать новые записи на основе КС, научность и достоверность которых подтверждались авторитетными источниками.

В первую очередь из словаря выводились «объектные» наименования: персоналии, родовые имена, заголовки организаций, заглавия произведений, географические наименования. КС на перечисленные наименования были выявлены в словаре и заменены АЗ, которые в БЗ соотносились с соответствующими полями 6 Блока анализа содержания формата BELMARC [1].

В процессе редактирования словаря «Неконтролируемые тематические термины» происходило разведение наименований объектов по видам, т. е. все понятия уже собирались в соответствующих словарях, и поисковые термины больше не дублировались в качестве КС и АЗ. Это позволяло пользователям при конкретном запросе работать с одним словарем, а не проводить поиск по разным словарям ЭК, чтобы выявить все возможные варианты.

Далее продолжалось редактирование КС, которые были составлены на тематический предмет. Они заменялись АЗ соответствующего содержательного наполнения или сочетанием нескольких дескрипторов, создавались и новые АЗ. При редактировании выполнялась и лексико-грамматическая обработка терминов, которые решено было оставить в составе словаря КС.

В 2022 г. было завершено редактирование словаря «Неконтролируемые тематические термины». На сегодняшний день он включает КС, которые прошли смысловую обработку и доработаны до принятой формы в соответствии с методическими решениями для ЭК и СЭК.

По окончании редактирования словаря КС была начата работа со словарями АЗ. В настоящее время уже завершена проверка словаря «Торговая марка (товарный знак)» и выполняется редактирование словарей «Тематический предмет» и «Географическое название». В ходе данной работы проверяется структура АЗ, проводится семантико-сопоставительный анализ иностранной и русской терминологии, в случае ошибок при синхронизации записей в локальном каталоге вручную обновляются АЗ. Обязательно контролируется нормативное представление дескрипторов в поисковых образах, при неверном употреблении терминов осуществляется переиндексирование БЗ. При этом из ПОД удаляются АЗ, значение которых изменилось и не соответствует содержанию документа.

В процессе работы со словарями ЭК библиотеки и сводного каталога выполняется и редактирование АЗ, которое осуществляется непосредственно в СЭК. Эта деятельность обеспечивает упорядочение логико-семантической структуры словарей, существенно влияет на качество смысловой обработки документов, дает четкую информацию по применению терминов, позволяет исключить субъективизм при выборе из словарей дескрипторов по определенной теме, что приводит к эффективному и резуль-

тативному выполнению поисковых запросов. В рамках редактирования АЗ исправляются ошибки в заголовках записей, устанавливаются иерархические и ассоциативные связи между терминами, вносятся изменения для корректировки смыслового наполнения – дорабатываются справочные примечания, добавляются отраслевые реляторы. В примечаниях уточняются сведения, которые конкретизируют использование АЗ: определение термина, варианты наименования, синонимы, в том числе условные [2].

В процессе редактирования АЗ приходится решать и проблемы синонимии в тех случаях, когда АЗ, схожие по семантическому наполнению, становятся избыточными для создания поискового образа конкретного документа. Такое редактирование способствует единообразию ПОД, так как наличие в словарях АЗ, совпадающих по содержанию, приводит к возможности использования при тематическом поиске разных терминов, что не способствует получению полной информации по запросу пользователей.

Поскольку процесс редактирования включает не только анализ и доработку АЗ, но и проверку поисковых образов, то в значительной степени исправление и замена терминов осуществляются вручную. Тем не менее для проведения корректировки терминов, особенно в большом массиве БЗ, нередко используются и программные средства. В связи с внедрением в ЦНБ НАН Беларуси комплекса информационно-технологических систем для автоматизации научных и научно-технических библиотек на основе облачных web-технологий (КИТС БИТ WEB) необходимы были новые программные инструменты, которые заменили бы программы, используемые при работе в автоматизированной библиотечно-информационной системе БИТ-2000и. Разработчиками КИТС БИТ WEB были составлены программы «Корректировка терминов в ЭК» и «Удаление прототипов», которые позволяют проводить актуализацию ЭК и программную корректировку АЗ и КС при редактировании каталога.

Программа «Корректировка терминов» позволяет одновременно заменить один термин (КС и АЗ) на три новых, что значительно упрощает выполнение сложных случаев редакции (рис. 1). Замена КС и АЗ осуществляется во всех БЗ, в которых встречается исключаемый термин.

Рис. 1. Пример замены КС на АЗ с использованием программы «Корректировка терминов»

При работе со словарем «Неконтролируемые тематические термины» в локальном каталоге остаются КС, исключенные из поискового образа и больше не связанные с БЗ. В таком случае используется программа «Удаление прототипов», которая автоматически формирует списки КС и прототипов АЗ (в соответствии с выбранным энтитетом), не имеющих связей с БЗ, и выполняет удаление этих терминов из каталога

(рис. 2). Использование такой программы позволяет уменьшать временные затраты на завершающем этапе редактирования и поддерживать корректность словарей.

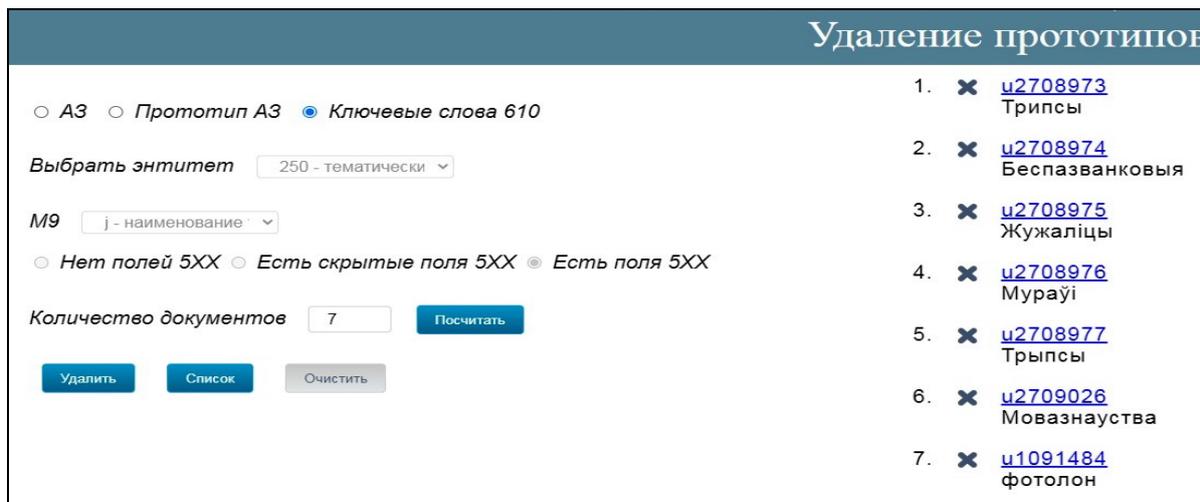


Рис. 2. Пример удаления из ЭК исключенных КС с использованием программы «Удаление прототипов»

Немаловажное значение для получения пользователями библиографической информации по определенным тематическим запросам имеет проведение поиска по классификационным индексам. При формировании поискового образа с помощью индексов УДК можно достаточно точно и глубоко обозначить смысловое содержание документа. Для достижения этой цели индексаторы ОЛЮ стараются использовать все возможности классификации, соблюдая принятые в ней правила индексирования и грамматики. Этапы индексирования включают анализ содержания документа, принятие и оформление классификационного решения, занесение классификационного индекса в БЗ в соответствии с правилами формата BELMARC.

УДК является стандартизированной классификационной системой, охватывающей все области знаний, которая позволяет группировать данные как в широком, так и узком тематическом диапазоне и обеспечивать обзорность фондов библиотеки. Для поддержания соответствия нынешнему уровню развития отраслей науки таблицы УДК регулярно перерабатываются: исключаются разделы, которые потеряли свою значимость, и вводятся новые, соответствующие актуальным информационным потребностям. Поправки, внесенные в таблицы классификации, отражаются в бюллетенях «Изменения и дополнения в УДК» [3]. Все изменения необходимо вносить в ЭК библиотеки, поэтому по мере необходимости и в зависимости от актуальности темы планируется и проводится редактирование разделов УДК. В процессе редактирования ЭК по разделам УДК осуществляются проверка БЗ и доработка поискового образа, внесение исправлений в индексы согласно актуальным таблицам УДК, что, безусловно, влияет на качество выполнения тематических запросов.

Особенность проведения редактирования разделов УДК в ЭК заключается в поиске и отборе БЗ, составленных на документы, связанные с определенными отраслевыми разделами и классами. До 2015 г. в ЦНБ НАН Беларуси проводилось редактирование систематического каталога, где карточки с описанием изданий собирались за конкретными тематическими разделителями, что обеспечивало наглядность раздела и удобство редактирования, в ходе которого выполнялась пересистематизация документов и в индексы УДК вносились исправления. По окончании редактирования все исправления переносились в записи ЭК.

После закрытия систематического каталога в 2015 г. разделы УДК редактируются только в ЭК. Выявление БЗ, относящихся к нужному разделу, осуществляется путем выполнения «профессионального поиска» (по индексам и булевскому поиску) и «расширенного поиска», при котором составляется как простой запрос по УДК, так и комбинированный по нескольким областям поиска.

При поиске по УДК в зависимости от выбора или крупного отраслевого раздела, или достаточно узкого тематического подраздела в поисковое поле классификационный индекс может вводиться как полностью, так и с усечением (с помощью «?»). Булевский запрос формируется в виде поискового выражения, состоящего из кода «udc» и поискового термина, в данном случае индекса УДК, соединенных через « = ».

Для проведения тематического поиска по индексам УДК пользователи также применяют «профессиональный поиск» и «расширенный поиск» в ЭК библиотеки.

Работа со словарями и редактирование разделов УДК влияют на качество ЭК, существенно повышают его поисковые возможности, улучшают выполнение информационных запросов. Редактирование ЭК позволяет уточнить или составить новый ПОД. Это дает возможность снять многозначность поискового образа, сделать его более информативным, максимально раскрывающим содержание документа, обеспечить пользователям корректное проведение в ЭК поиска по словарям и с помощью индексов УДК, получить удовлетворительный результат на свой тематический запрос.

Список литературы

1. Астапович, Л. Л. Ведение словаря ключевых слов электронного каталога Центральной научной библиотеки НАН Беларуси / Л. Л. Астапович // Развитие информатизации и государственной системы научно-технической информации (РИНТИ-2021) : доклады XX Междунар. науч.-техн. конф., Минск, 18 ноября 2021 г. – Минск : ОИПИ НАН Беларуси, 2021. – С. 257–261.

2. Астапович, Л. Л. Авторитетный контроль при формировании поискового образа документов в ЦНБ НАН Беларуси / Л. Л. Астапович // Развитие информатизации и государственной системы научно-технической информации (РИНТИ-2022) : доклады XXI Междунар. науч.-техн. конф., Минск, 17 ноября 2022 г. – Минск : ОИПИ НАН Беларуси, 2022. – С. 293–297.

3. Астапович, Л. Л. Актуализация Универсальной десятичной классификации: анализ изменений и дополнений к классификационной системе / Л. Л. Астапович // Библиотека в XXI веке. Социальная миссия : материалы XIII Междунар. науч.-практ. конф. молодых ученых и специалистов, Минск, 24–25 марта 2022 г. – Минск : Ковчег, 2022. – С. 238–244.