

ИНТЕГРАЦИЯ ИНФОРМАЦИОННЫХ РЕСУРСОВ БИБЛИОТЕК БЕЛАРУСИ

INTEGRATION OF INFORMATION RESOURCES OF BELARUSIAN LIBRARIES

Шереметьева Алла Анатольевна – начальник информационного центра, Национальная библиотека Беларуси (Беларусь), e-mail: sheremeteva@nlb.by

Бричковский Вячеслав Иванович, заведующий сектором, кандидат технических наук, Национальная библиотека Беларуси (Беларусь), e-mail: v_britch@nlb.by

Sheremeteva Alla – head of the Information Center, National Library of Belarus (Belarus), e-mail: sheremeteva@nlb.by

Britchkovski Viatcheslav – head of the sector, candidate of technical sciences, National Library of Belarus (Belarus), e-mail: v_britch@nlb.by

Аннотация: Рассмотрены проблемы и перспективы развития сводного электронного каталога библиотек Беларуси. Описаны базовые подходы к формированию платформы сводного каталога библиотек: интегрированная система идентификации ресурсов, единая система поиска и доступа к ресурсам, федеративная система авторизации и аутентификации пользователей на основе ID карты.

Abstract: The problems and prospects of the development of the unified electronic catalog of the libraries of Belarus are considered. The basic approaches to the creation of the platform of the unified catalog of libraries have been proposed: an integrated system for identifying resources, a unified system for searching and accessing resources, a federated system of authorization and authentication users based on an ID card.

Ключевые слова: сводный электронный каталог, интеграция информационных ресурсов, discovery сервисы, DOI, ROR, ORCID.

Keywords: unified electronic catalog, integration of information resources, discovery services, DOI, ROR, ORCID.

В библиотечной практике первым направлением интеграции информационных ресурсов, до сих пор успешно функционирующим и активно развивающимся, являются сводные электронные каталоги библиотек. Коллаборация библиотек в части каталогизации информационных ресурсов способствует:

- сокращению производственных и финансовых затрат;
- обеспечению унификации описания ресурсов;
- повышению качества метаданных.

Практически в каждой стране есть примеры таких корпоративных проектов и ресурсов. В Бельгии – UniCat, во Франции – SUDOC, в Швеции – LIBRIS, в Польше – NUKAT и др.

В Сводном электронном каталоге (СЭК) библиотек Беларуси представлены фонды 4-х крупнейших библиотек страны, общим объемом более 8 млн записей. На базе СЭК функционирует система корпоративной каталогизации изданий [1]. Свыше 150 библиотек ежегодно экспортируют в свои локальные каталоги около 60 тыс. качественных записей из СЭК. В среднем каждая новая запись каталога заимствуется трижды.

Для нормирования и контроля лексического состава информационно-поискового языка, расширения возможностей связывания имен и объектов метаданных в СЭК используется технология авторитетных записей, с применением которой создается национальное пространство имен авторов и организаций,

а также универсальный дескрипторный словарь. Это позволяет:

- однозначно идентифицировать объект описания, выделить его среди других объектов в описательных метаданных, имеющих такое же или сходное имя;
- связывать разные имена одного объекта;
- устанавливать полезные для информационного поиска связи разных объектов;
- обеспечивать лингвистическую совместимость библиографических записей (БЗ), подготовленных разными библиотеками.

С 2016 года активно формируется сеть региональных сводных каталогов (РСЭК) библиотек Беларуси. Созданы шесть РСЭК. В их состав в настоящее время вошли 48 библиотечных учреждений культуры и образования. Эти каталоги являются агрегаторами метаданных библиотек области, информационно-поисковыми системами краеведческих документов, а также выполняют функции региональных систем корпоративной каталогизации документов [2].

Оператором СЭК и всех РСЭК является Национальная библиотека Беларуси. Она координирует работу по формированию каталогов, ведет методическое сопровождение каталогов, контролирует качество записей, сопровождает программно-технический комплекс.

Однако возможности текущей версии СЭК в создании национального библиографического контента ограничены. Не все библиотеки нашей страны могут работать с метаданными СЭК. Не решены вопросы обмена данными между библиотеками страны, которые используют при каталогизации форматы, отличные от BELMARC. Значительные изменения происходят и в самой информационной среде. Бурное развитие сетевых технологий способствует значительному увеличению объемов информационных ресурсов и появлению «информационного шума». Проблема осложняется и доступностью полных текстов документов для пользователей. Они могут быть представлены как на платных платформах издательств и агрегаторов, так и в виде электронных ресурсов открытого доступа.

Эти и ряд других факторов требуют совершенствования системы СЭК.

В основе дорожной карты развития СЭК были определены три базовых блока, реализация которых позволит осуществить эффективную интеграцию и использование информационных ресурсов в масштабе страны.

Эта «триада» включает:

- интегрированную систему идентификации ресурсов;
- единую систему поиска и доступа к ресурсам;
- федеративную систему авторизации и аутентификации на основе ID карты.

Интегрированная система идентификации ресурсов.

Она должна быть основой развития ресурсов СЭК и обеспечить создание национального пространства публикуемых документов, авторов публикаций и организаций.

Такая информация уже частично представлена в базе данных (БД) авторитетных записей СЭК. Однако в настоящее время необходима модификация этой БД и ее интеграция с мировыми системами идентификации, которые уже широко используются в системе научных коммуникаций.

Для идентификации авторов наиболее оптимальным решением является использование ORCID. Это алфавитно-цифровой код, который позволяет ученому объединить в одном профайле данные обо всех своих публикациях, аффилиациях, исследованиях и научных процессах, в которых он участвует. ORCID визуально и технически связывает любые научные процессы с их авторами, помогает поддерживать список публикаций в актуальном виде и предоставляет возможность производить интеграцию с другими информационными сервисами. ORCID получил широкое распространение, поддерживается некоммерческой организацией и обладает возможностями интеграции с другими системами авторских профилей.

При реализации системы идентификации организаций следует ориентироваться на проект Реестр исследовательских организаций (ROR) (<https://ror.org/>). Он обеспечивает уникальную и постоянную идентификацию организаций и связь с соответствующими метаданными, описывающими организацию. Проект поддерживает бесплатный открытый API (англ. application programming interface - программный интерфейс приложения) для управления записями реестра и их интеграции с другими системами идентификаторов типа ORCID и DOI.

В условиях глобализации и бурного развития Интернет актуальной становится проблема создания единой международной системы регистрации электронных информационных ресурсов.

Для однозначной идентификации документов национального информационного пространства целесообразно использовать систему цифровых идентификаторов DOI. Она предназначена для идентификации информационных объектов в цифровой среде и в настоящее время широко используется в сфере научных коммуникаций. В основе технологии DOI лежит принцип, согласно которому однажды ассоциированный с контентом индекс должен оставаться

неизменным и связанным именно с этим материалом. Метаданные объектов и их адреса могут меняться, но сами идентификаторы всегда остаются постоянными.

В рамках международной системы CrossRef на основе технологии DOI уже создан единый реестр научных информационных материалов и связанных с ними метаданных. Кроме того, в настоящее время поставщики ведущих баз данных не признают электронные публикации без DOI, так как без него сложно предоставить устойчивый доступ к ресурсам средствами Интернет, а это отрицательно сказывается на индексировании и цитируемости публикаций.

Единая система поиска и доступа к ресурсам.

Для развития СЭК библиотек Беларуси приоритетными остаются задачи более полной индексации информационных ресурсов белорусских библиотек и их интеграции с лицензионными ресурсами издательств и агрегаторов, коллекциями ресурсов открытого доступа.

В настоящее время при создании каталогов библиотек активно используется технология построения объединенного поискового индекса (discovery-сервис). Принцип работы discovery-сервиса сводится к созданию единого поискового индекса на различные виды ресурсов, доступных как в фонде библиотеки, так и за ее пределами [3]. Поиск осуществляется на основе этого индекса. В ответ на запрос пользователь получает результаты в виде метаданных со ссылками на полные тексты документов, доступных из фондов библиотеки, баз данных закупленных лицензионных ресурсов, коллекций ресурсов открытого доступа. Путь к полному тексту документа формируется в результате работы модуля Link Resolver (определителя ссылок).

Указанная система позиционируется как «каталог следующего поколения», цель которого – обеспечить более совершенный вариант онлайн-каталога публичного доступа (OPAC). Она предоставляет улучшенный интерфейс для конечных пользователей при построении и выполнении запросов и обеспечивает повышение релевантности поиска, фасетную навигацию и другие преимущества при поиске и работе с онлайн-ресурсами.

Такое решение может стать частью облачной национальной платформы для управления как печатными, так и электронными ресурсами, обладающей сервисной архитектурой (SOA).

В НББ опыт использования технологии построения объединенного поискового индекса уже есть. В конце 2018 г. в НББ была создана новая версия портала библиотеки и Электронного каталога (ЭК) [4]. За основу разработки каталога взято ПО VuFind и Solr. В процессе работы были учтены основные вопросы сопряжения АБИС НББ и VuFind. В их числе: создание метаданных в формате BELMARC, использование в метаданных авторитетного контроля, технология иерархически связанных записей для описания многотомных документов и аналитических записей. Для подключения к поиску в каталоге подписных, лицензионных электронных ресурсов в каталог был интегрирован сервис discovery компании EBSCO. Читатели НББ получили возможность через каталог работать с лицензионными БД 7 контент-провайдеров ресурсов.

Дополнительно был подключен плагин Unpaywall, который обеспечивает поиск документов открытого доступа на основе DOI.

Федеративная система авторизации и аутентификации

Для организации удаленного доступа к ЭИР и сервисам сводного каталога предлагается внедрение системы федеративной идентификации, аутентификации и авторизации.

Система предназначена для решения следующих задач:

- обеспечение персонифицированного доступа к ЭИР в любой точке мира с устройства, подключенного к Интернет;
- популяризация авторитетных ЭИР в научно-образовательной среде;
- организация эффективного информационного взаимодействия обслуживающего персонала библиотеки с правообладателями, поставщиками и потребителями ЭИР;
- учет и анализ статистики использования ЭИР.

С помощью этой системы пользователи из организаций-партнеров смогут безопасно и конфиденциально получать услуги, используя только аккаунт (учетную запись) в своей организации при помощи технологии сквозной аутентификации (Single Sign-On или SSO). Таким образом, для доступа к ресурсам конечному пользователю не нужно сохранять какие-либо дополнительные или особенные логины и пароли. Если его организация подключена к системе, то из любой точки он может получить доступ к тому набору ресурсов, на которые подписана его организация.

Система обеспечит надежный обмен информацией, связанной с идентификацией, аутентификацией и авторизацией пользователей. В основе такой системы лежит идентификационная федерация. Это ассоциация организаций, целью которой является сотрудничество на основе обмена учётными данными. Деятельность федерации направлена на упрощение внедрения сервисов и служб. В состав федерации входят провайдеры учётных данных, предоставляющие информацию о пользователях и провайдеры сервисов, использующие эти данные для оказания услуг. При этом авторизация пользователя происходит исключительно на серверах своей организации. В качестве источника записей о пользователях при этом могут выступать различные базы данных, используемые в организации: базы пользователей контролеров доменов (Active Directories, LDAP) и другие базы данных, в частности база данных читателей на основе ID-карты.

Наиболее перспективным подходом к реализации систем федеративной идентификации является применение технологии SAML. Одной из ключевых причин, обусловивших активное развитие и широкое применение SAML, является продуманная и развитая архитектура, позволяющая строить на основе этой технологии самый широкий спектр пользовательских решений, ориентированных на дальнейшую интеграцию с государственными информационными системами на основе ID-карты.

Практическая реализация предложенных подходов позволит создать платформу для формирования сводного электронного каталога библиотек

Беларуси с возможностью интегрированного поиска и безопасного доступа как к национальным информационным ресурсам, так и многочисленным лицензионным ресурсам ведущих мировых производителей.

СПИСОК ЛИТЕРАТУРЫ

1. Шереметьева, А. А. Сводный электронный каталог библиотек Беларуси как интегрирующая система метаданных / А. А. Шереметьева // Материалы IV Международного конгресса «Библиотека как феномен культуры». Информационные ресурсы библиотек в образовательной, научной и социокультурной среде, Минск, 4–6 окт. 2016 г. – Минск, 2016. – С. 183–188.
2. Шереметьева, А. А. Региональные сводные электронные каталоги как инструмент корпоративного взаимодействия библиотек Беларуси / Алла Шереметьева // Бібл. свет. – 2018. – № 1. – С. 29–31.
3. Бричковский, В. И. Реализация систем интегрированного поиска в библиотеках с использованием программного обеспечения с открытым исходным кодом / В. И. Бричковский, А. А. Шереметьева // Развитие информатизации и государственной системы научно-технической информации : РИНТИ-2020 : XIX Междунар. конф., Минск, 19 нояб. 2020 г. : докл. – Минск, 2020. – С. 310–314.
4. Шереметьева А. А. Каталоги библиотек нового поколения и сервисы Discovery: опыт Национальной библиотеки Беларуси / Алла Шереметьева, Маргарита Никифорова, Зоя Петровская // Бібл. свет. – 2019. – № 2. – С. 3–6.