

УДК 025.5/6:024.5

В.И. Бричковский¹, В.И. Комик², Н.И. Листопад³, В.В. Мицкевич⁴

ПЕРСПЕКТИВЫ ИНТЕГРАЦИИ НАУЧНО-ОБРАЗОВАТЕЛЬНЫХ ИНФОРМАЦИОННЫХ РЕСУРСОВ ОТКРЫТОГО ДОСТУПА

Анализируется современное состояние генерации, использования и распространения электронных информационных ресурсов. Рассматриваются вопросы интеграции информационных ресурсов и сервисов на основе применения передовых технологий работы с метаданными. Предлагается подход, направленный на создание единого информационного пространства научно-образовательных ресурсов открытого доступа.

Во всем мире активно развивается новая модель публикаций в научно-образовательной сфере, связанная с созданием открытых архивов информации. Успех этой модели базируется на гармоничном сочетании интересов всех участников коммуникационного процесса. Такая модель дает возможность пользователям получить доступ к информации, библиотекам – обеспечить этот доступ, издательствам – предоставлять информацию в соответствии с запросами ее потребителей [1].

Концептуально суть инициативы открытых архивов состоит в том, чтобы противопоставить парадигме «знание – товар» некоммерческий принцип открытого доступа (ОД) к результатам научной и образовательной деятельности для самого широкого круга пользователей в сетевом режиме [2]. Автор может разместить копии своих научных статей в институциональном репозитории ОД или опубликовать свою работу в одном из журналов, поддерживающих в той или иной форме инициативу ОД. При этом в обоих случаях для таких статей будет обеспечена гораздо более широкая доступность, чем для статей, размещенных в базах данных (БД) журналов с платной подпиской [3, 4].

Как следствие пользователям при поиске информации приходится работать с множеством разнообразных БД, сильно различающихся между собой по структуре, тематической направленности и полноте представления. Многие организации, сталкивающиеся с необходимостью предоставления доступа к своим информационным ресурсам (ИР), разрабатывают специальные системы для работы с ними. Пользователям в таком случае предоставляется специализированный интерфейс доступа к информации, содержащейся в поддерживаемой БД, и при работе с такими системами приходится сталкиваться с различными подходами к поиску и навигации. Зачастую большая часть времени тратится на изучение интерфейса, предоставляемого системой, и его возможностей, а не на работу с информацией.

Эти причины обуславливают целесообразность применения подходов, ориентированных на универсальный интерфейс доступа к БД ИР, различным по структуре и содержанию.

Анализ современных тенденций свидетельствует о стремлении к виртуальному объединению информационных и вычислительных ресурсов на уровне предоставления доступа к ним [5].

Решения, представленные в данной статье, ориентированы на процессы создания, поддержки, интеграции и распространения электронных ИР открытого доступа в научно-образовательной сфере на базе межвузовского портала ОД (<http://metaport.unibel.by>) [6].

Текущее состояние работы с ИР открытого доступа характеризуется следующими особенностями:

- увеличивающейся сложностью решаемых задач при одновременной необходимости сокращения времени обработки информации;
- достаточно большим объемом информации;
- различными схемами метаданных;
- разнообразием видов используемых ИР.

Эти особенности обуславливают всевозрастающие требования по повышению эффективности процессов обработки и использования ИР.

Предлагаемый подход позволит автоматизировать часть процессов по подготовке и размещению метаданных на межвузовском портале ОД, а также процессы обслуживания и администрирования портала, включая процессы поиска необходимой информации.

Данные процессы направлены на решение следующих задач:

- автоматизации процессов формирования и обработки ресурсов ОД в научно-образовательной сфере;
- автоматизации процессов сбора метаданных из архивов ОД и электронных библиотек вузов;
- создания «единой точки входа» (навигационной системы) на основе веб-технологий для получения комплексной информации для обеспечения научно-образовательных процессов;
- создания интегрированного хранилища метаданных;
- реализации процедур по преобразованию схем и форматов метаданных;
- обеспечения эффективного доступа к научно-образовательной информации на базе веб-технологий;
- администрирования прав доступа различных категорий пользователей к функциям межвузовского портала ОД.

При выборе структурных решений учитывалось, что система должна позволять хранить необходимые объемы информации, а также поддерживать гибкий механизм своевременной актуализации контента и базовый набор сервисов для работы с информацией.

Структура представления ИР и пользовательские интерфейсы по доступу к данным ресурсам и сервисам должны быть интуитивно понятны широкому кругу пользователей Интернета.

При выборе структуры системы необходимо, чтобы она обеспечивала реализацию следующих основных сервисов:

- хранения метаданных;
- сбора метаданных из информационных источников;
- преобразования схем метаданных;
- управления представлением информации на портале;
- поиска информации по атрибутам метаданных;
- администрирования пользователей системы.

Первый сервис управляет хранением метаданных и атрибутов информационного источника (data-provider). Второй сервис обеспечивает сбор метаданных из информационных источников хранения данных. В соответствии с этим необходимо применение различных гарантирующих механизмов, таких как обнаружение сбоев при сборе метаданных и смягчение эффектов этих событий.

Внешний или внутренний источник комплектования получает подтверждение того, что собранные данные инкорпорированы, скомплектованы и преобразованы (в случае необходимости) из информации, переданной на рассмотрение, в форму, подходящую для хранения и управления внутри хранилища системы, извлечения и (или) создания описательных метаданных, которые поддерживаются поиском.

Функции, связанные с сервисом доступа, заключаются в управлении процессами взаимодействия с пользователями информации.

Служба доступа обеспечивает поддержку пользователя информации и обрабатывает запросы, предоставляет формирование результирующих списков по ним и координирует восстановление и доставку информации, получает запрошенные сведения и выполняет любые необходимые преобразования (возможные трансформации включают изменения формата на более удобный для распространения и маскируют определенные типы метаданных). Служба доступа также отвечает за реализацию механизмов безопасности или механизмов контроля доступа к контенту.

Сервис администрирования отвечает за управление учетными записями и правами пользователей, а также за координацию деятельности других сервисов.

Таким образом, обеспечивается:

- хранение всех типов метаданных в интегрированном хранилище;
- минимизация ручного труда при сборе и индексации метаданных;

- доступность информации с любого клиентского компьютера, на котором установлен только веб-браузер;
- возможность нахождения ресурса по его метаданным;
- доступность функций портала только для пользователей с соответствующими правами доступа.

Для этого предлагается использовать протоколы, позволяющие провайдерам информации снабжать свои данные набором метаданных, которые затем собираются и обрабатываются сервисным провайдером, создающим, по сути дела, новый ИР. В этом случае сервисному провайдеру нет необходимости поддерживать интерактивную связь между конечным пользователем и выбранным им источником информации. Он собирает и индексирует метаданные заранее. Пользователь получает возможность работать с множеством ресурсов через один поисковый интерфейс. При этом метаданные могут включать ссылки на полнотекстовые и мультимедийные документы в соответствующем архиве.

Для построения системы было предложено использовать модель со смешанным хранением ресурсов и протоколом обмена OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting).

Такой подход, ориентированный на технологию комбинированного хранения данных, имеет ряд преимуществ. Во-первых, поиск по метаданным идет очень быстро, так как все они хранятся в одном месте на одном сервере. Нет необходимости опрашивать постоянно все серверы, так как это может приводить к дополнительным временным задержкам с учетом того, что ширина полосы телекоммуникационных каналов в библиотечных корпоративных сетях достаточно узка. Во-вторых, пользователь при организации поиска может использовать произвольный набор метаданных.

Сбор информации из ресурсных центров осуществляется по протоколу OAI-PMH, расширенному для решения задач целенаправленного отбора метаданных.

Есть две альтернативы: точно настраивать избирательный отбор для конкретного контент-провайдера или проводить внутреннюю обработку собранных метаданных на портале. Предлагается ввести расширение OAI-PMH для селективного сбора ресурсов с учетом критериев, которым должна удовлетворять собираемая информация [4].

В результате сборщик неявно передает критерии провайдеру информации и тот способен динамически сформировать и вернуть набор, в который включены все документы, отвечающие заданным критериям. При этом важно отметить, что расширение отвечает спецификации протокола OAI-PMH.

В настоящее время в Главном информационно-аналитическом центре Министерства образования Республики Беларусь реализован пилотный проект по созданию межвузовского портала метаданных ресурсов ОД.

Предложенные решения будут способствовать появлению научно-образовательных информационных пространств ОД, основанных на федеративных принципах. Новым качеством, возникающим в результате создания такого портала, является предоставление пользователям единых интерфейсов доступа к интегрированным разнородным ресурсам на основе сбора, обработки и хранения метаданных.

Список литературы

1. Campaign for free access // RTD Info (Magazine on European Research). – 2005, november (special issue). – P. 11–12.

2. Возможности использования информационных ресурсов открытого доступа в сфере образования / Н.И. Листопад [и др.] // Доклады IX Междунар. конф. «Развитие информатизации и системы научно-технической информации (РИНТИ-2010)», Минск, 18 нояб. 2010 г. – Минск : ОИПИ НАН Беларуси, 2010. – С. 190–192.

3. Просер, Д. Очередная революция в информации – могут ли репозитории и самоархивирование изменить систему научного общения / Д. Просер // Научные и технические библиотеки. – 2003. – № 11. – С. 32–41.

4. Московкин, В.М. Два пути «интернетизации» научных исследований в условиях глобализации / В.М. Московкин, Е.М. Савченко // Научно-техническая информация. Сер. 1. Организация и методика информационной работы. – 2008. – № 1. – С. 16–18.

5. Сенько, А.М. Система виртуальной интеграции разнородных баз данных / А.М. Сенько, М.М. Якшин // Технологии информационного общества – Интернет и современное общество : труды VII Всерос. объединенной конф., Санкт-Петербург, 10–12 нояб. 2004 г. – СПб. : Изд-во филолог. ф-та СПбГУ, 2004. – С. 142–143.

6. Бричковский, В.И. Применение современных подходов к интеграции электронных информационных ресурсов в сфере культуры / В.И. Бричковский // Материалы VI науч.-практич. конф. «Культура, наука, творчество», Минск, 10–12 мая 2012 г. – Минск : БГУКИ, 2012. – С. 337–341.

Поступила 10.11.2012

¹Национальная библиотека Беларуси,
Минск, пр. Независимости, 116
e-mail: v_britch@nlb.by

²Главный информационно-аналитический центр
Министерства образования Республики Беларусь,
Минск, ул. Захарова, 59
e-mail: giac@giac.unibel.by
komik@unibel.by

³Белорусский государственный университет
информатики и радиоэлектроники,
Минск, ул. П. Бровки, 6
e-mail: syst@bsuir.by

V.I. Brichkovskij, V.I. Komik, N.I. Listopad, V.V. Mitskevich

INTEGRATION OF OPEN ACCESS SCIENTIFIC AND EDUCATIONAL INFORMATION RESOURCES

The current state of the generation, use and dissemination of electronic information resources is analyzed. The integration of information resources and services through the application of advanced technologies of metadata have been considered. An approach aimed at creating a single information space of research and educational resources of open access is offered.