

Электронные ресурсы в библиотеке Казанского федерального университета (От электронных каталогов к системам поиска и доставки ресурсов)

Струков Евгений Николаевич, директор, Научная библиотека им. Н. И. Лобачевского Казанского федерального университета, e.strukov@list.ru

В докладе приводится краткий анализ современных тенденций развития библиотечных информационно-коммуникационных систем на основе требований, предъявляемых пользователями к реализации поисковых возможностей. В качестве одного из вариантов open source решения предлагается front-end система поиска и интеграции электронных библиотечных ресурсов Vu Find, на основе которой модернизирован портал библиотеки КФУ. Данная система также в последствии будет интегрирована с Printgroove JT Suite 5.0 для реализации «Печати по требованию». В дополнению к упомянутым выше решениям, описана реализация модуля Имидж-каталога с возможностью контекстного поиска.

Общеизвестно, что за последние десятилетия произошли кардинальные изменения статуса библиотеки. На протяжении длительного времени книжные фонды были одним из немногих источников достоверной информации, и пользователи были обязаны следовать библиотечной практике и стандартам для успешного удовлетворения информационных потребностей. Сегодня ресурсы имеются в изобилии, а поисковые стратегии, привычки и предпочтения пользователей формируются Интернетом и новыми лидерами информационного рынка (Google и пр.). Результаты многочисленных зарубежных исследований пользователей академических библиотек свидетельствуют о снижении роли библиотеки как отправной точки или шлюза для проведения исследований. Аналогичные тенденции прослеживаются и в России.

В частности, анкетирование пользователей библиотеки КФУ дало следующие результаты: 65,5% опрошенных аспирантов КФУ 1-го года обучения начинают поиск информации по теме исследования в Интернете и лишь 38,7% с электронного каталога библиотеки. В сети ищут нужную информацию с помощью поисковых сервисов - 87% респондентов, обращаются к электронным ресурсам (базы данных, электронные журналы) - 61,20%, сайтам библиотек - 31,10%, и просматривают тематические сайты - 19,3%. 64% опрошенных считают, что самостоятельно могут подобрать всю необходимую им информацию. Около половины респондентов на вопрос: «Каких сервисов, по Вашему мнению, не хватает в библиотеке в настоящее время?» ответили, что не хватает электронных вариантов книг и журналов, удобного «поисковика», сервиса оповещения о новинках по электронной почте и др.

Обобщая результаты различных исследований, можно выделить, отнюдь не полный, список причин, обусловивших сложившуюся ситуацию и набор претензий пользователей к библиотечным поисковым системам:

- Отсутствие простых и прозрачных путей от поиска до извлечения информационных источников, причем для многих пользователей решающее значение имеет доступ к полному тексту.
- Не соответствие результатов поиска ожиданиям и низкая оценка качества данных, формируемая пользователем, главным образом, на собственном опыте и зависящая от соответствия поисковых и навигационных интерфейсов библиотеки поисковым системам Интернета и популярным веб-сайтам.
- Низкая оценка релевантности результатов поискового запроса, с большой долей вероятности связанная с тем, что пользователи не различают каталог, базы данных, цифровые коллекции, а также бесплатные ресурсы научного содержания, и ряд других.

Совершенствование системы доступа к информационным ресурсам – традиционный вектор развития информационных технологий в университетской библиотеке, он важен сейчас для сохранения решающей роли библиотечных сервисов в процессах обучения и научных исследованиях. Эта тенденция становится еще более значимой в связи с существенным повышением доли электронных документов различного происхождения (сканированные изображения редких изданий и рукописей, полнотекстовые документы, права на которые принадлежат вузу и проч.) в фондах библиотек и формировании на их основе IT-решений, которые можно объединить под названием «полнотекстовая электронная библиотека».

Результаты анализа поведенческих стратегий и претензий пользователей к существующим каталогам нашли отражение в дизайне и сервисах нового поколения, называемых *Webscale Discovery Service (поиск и извлечение ресурсов в масштабе Web)*. Коммерческие и программные решения с открытым кодом в той или иной мере реализуют базовый набор поисковых возможностей, навигации и обработки результатов запроса:

- простой интуитивно понятный поисковый интерфейс, предоставляющий возможность проводить поиск по ключевым словам с использованием простого поискового окна и быстрый переход к расширенному поиску для составления более сложных запросов;
- широкие возможности фасетного поиска и навигации в больших массивах результатов поиска;
- ранжирование результатов поиска в соответствии с релевантностью запросу или предпочтениям пользователя;
- автозаполнение при формировании запроса и проверка орфографических ошибок;
- персонализация настроек.

Кроме того, предоставляются возможности использования социально ориентированных сервисов в стиле Web 2.0, дополнение традиционной каталожной записи данными из внешних источников и ряд других сервисов, отсутствующих ранее в электронных каталогах.

В основе реализации Webscale Discovery Service лежит всеобъемлющий централизованный индекс (в идеале, до уровня статьи), схема которого нормализована по типам контента.

Индекс создается путем сбора контента из:

- локальных ресурсов библиотеки;
- ресурсов издателей и агрегаторов, тех, что позволяют получить доступ к их метаданным и / или полному содержанию текста для индексирования.

Доступ / доставка информации (ресурса) пользователю обеспечивается за счет:

- включения цифровых и печатных ресурсов библиотеки на уровне статей (или документов);
- взаимодействия с сервером ссылок (link resolver) подключаемой библиотеки, чтобы обеспечить доступ на уровне статьи или службу доставки документов;
- интеграция с функциями библиотечного каталога для заказа печатных изданий.

Дизайн и сервисы библиотечных поисковых систем, в центре которых находится электронный каталог, должны трансформироваться и быть не хуже сходных сервисов, предлагаемых иными информационными поставщиками.

Проект электронной библиотеки в Казанском университете реализуется уже несколько лет силами сотрудников Центра цифровых технологий и сервисов Научной библиотеки им. Н.И. Лобачевского (рук. Центра В.Ю. Кузьмина).

В результате нескольких итераций мы пришли к следующей структуре программно-технических решений.

ЭБ КФУ базируется на двухуровневой модели front-end/back-end, базовым элементом (back-end решением) которой является электронный каталог (ЭК) АБИС Руслан. ЭК обеспечивает систематизацию документов, хранимых в цифровом формате, таких как коллекции сканированных изданий из редкого фонда, фотоархивы, электронные версии изданий авторов-сотрудников КФУ и т.п. Записи ЭК содержат необходимые метаданные и ссылки на файлы, доступ к которым возможен стандартными средствами Руслана. До определенного момента такая схема доступа казалась вполне достаточной, хотя и не отвечала современным требованиям, предъявляемым к поисковым сервисам в Интернет.

Необходимо отметить, что большинство библиотек крупных университетов России помимо собственно обеспечения своих посетителей книгами и другими печатными изданиями, как правило, предоставляют доступ к различным внешним электронным ресурсам, как зарубежным (Elsevier, ProQuest и проч.), так и отечественным (eLibrary, ЭБС), формируют собственные коллекции цифровых документов. Каждый из поставщиков упомянутых ресурсов имеет свой интерфейс поисковых запросов. Например, Казанский университет имеет доступ более, чем к 50 информационным источникам. Естественно возникает желание по возможности организовать поиск по этим БД в одном запросе, т.е. опять возни-

кает проблема создания интегрированного поискового интерфейса пользователя.

Сравнительно недавно появилась еще одна проблема - невозможность индексирования внешними поисковыми системами (например, Google Scholar) электронных ресурсов, доступ к которым осуществляется через ЭК АБИС. Необходимость такого индексирования связана с критериями отбора и продвижения в международных рейтингах вузов, таких как QS.

Таким образом, мы пришли к необходимости, с одной стороны, улучшить качество поискового сервиса, предоставляемого на сайте библиотеки, а с другой решить задачу интеграции всех(!) электронных ресурсов университета, по возможности, одним front-end решением.

Безусловно, данная задача не является чем-то принципиально новым, существует спектр программных продуктов, реализующих технологию Webscale Discovery. Достаточно вспомнить Serials Solutions' Summon, Ex Libris's Primo, EBSCO's Discovery Service и др. В основе программных продуктов этого класса лежит использование современных поисковых платформ (в частности, Apache Solr) для построения единого поискового индекса локальных библиотечных фондов, институциональных репозитариев и цифровых коллекций. Перечисленные программные комплексы являются коммерческими проектами и достаточно дорогими. Однако есть и альтернативные решения на той же основе, одним из таких программных продуктов является портал для доступа к ресурсам библиотеки – VuFind. Разработка была начата в Университете Вилланова США и продолжается сейчас в рамках Open source проекта.

VuFind – программное обеспечение с открытым исходным кодом, позволяющее реализовать современный интерфейс пользователя для работы с информационными ресурсами библиотеки. Программный инструментарий, не является полнофункциональной АБИС, но взаимодействует с ней в режиме реального времени для извлечения данных. В VuFind реализована связка с некоторыми популярными в Америке и на Западе библиотечными системами (драйвер ILS). Основное назначение VuFind – настраиваемый веб-интерфейс для ключевых библиотечных функций: информационного поиска и библиотечного обслуживания.

VuFind использует библиотеку SolrMARC для создания индекса по записям из библиотечных каталогов, а также предоставляет возможность проиндексировать полный текст. На основе построенного индекса реализуется гибкий механизм поиска по метаданным и полному тексту, а также простой интуитивно понятный интерфейс просмотра, который помогает изучить главные тенденции в коллекции или просто просмотреть основные элементы в алфавитном порядке.

Ключевые особенности VuFind:

- Фасетизация при выдаче результатов запроса - поиск и настраиваемые ограничения по множеству категорий.
- Формирование рекомендаций - выдача похожих документов по задаваемому алгоритму подобия, в том числе с использованием внешних источников.

- Сортировка результатов по релевантности и ряду других настраиваемых критериев. Постоянные ссылки на записи и возможность сохранять запросы и результаты поиска. Управление цитированием.
- Автозаполнение при формировании запроса.
- Поддержка возможностей веб 2.0.
- Многоязычный интерфейс.

В системе реализован личный кабинет пользователя. Каждый пользователь может зарегистрироваться, сохранять результаты поиска, добавлять ресурсы в избранное. Есть возможность просмотра всего каталога по различным категориям, отображаются последние поступления.

В настоящий момент в библиотеке КФУ с помощью этой системы проиндексирован каталог книг и загружен каталог электронных журналов и книг из подписки университета, включая платную подписку на зарубежные издательские БД и отечественные ЭБС – сегодня это «Книгофонд». Написан драйвер для связи с АБИС Руслан и реализован модуль «Заказ книг». Информация доступна читателю в его личном кабинете. Пока остается открытым вопрос построения иерархических связей между библиографическими записями (периодические издания). Еще одна проблема, с которой пришлось столкнуться: индексация записей для полнотекстового поиска. Эта проблема, как и ряд некоторых мелких ошибок в исходном коде, была решена с помощью обновления системы VuFind до последней версии (VuFind 2.0RC1, released 26.04.2013).

Следующий этап работ по развитию пользовательского интерфейса библиотеки КФУ заключается в разработке связки Vu Find с программным комплексом Printgroove JT Suite 5.0, который реализует поддержку технологии «Print-on-Demand» (Печать по требованию) на основе цифрового типографского комплекса на платформе печатного оборудования Konica-Minolta и финишного цикла других производителей. Контентная база формируется из учебно-методических материалов, созданных профессорско-преподавательским составом университета. В состав JT Suite входит подсистема управления печатным процессом, управления очередями заданий, Web-интерфейс клиентской части и многое другое. Система позволяет распределять поток заданий на несколько машин, реализует изготовление печатной продукции по заранее заготовленным шаблонам, печать переменных данных и проч.

Такая связка позволит любому пользователю после выполнения поискового запроса в ЭБ просмотреть электронную версию документа и, при необходимости, заказать твердую копию документа, получить которую он сможет после подтверждения оплаты заказа в одном из подразделений по обслуживанию читателей Научной библиотеки им. Н.И. Лобачевского.

Цифровой типографский комплекс позволяет изготавливать печатную продукцию широкого спектра, благодаря наличию в его составе черно-белой и полноцветной печатных машин, финишной линии твердого и мягкого переплета. Его возможности позволяют издавать не только рядовую книжную продукцию, но и достаточно сложные дизайнерские альбомы и журналы. В перспективе, на его основе будут выполняться заказные тиражи научных журналов, издаваемых КФУ (сегодня их насчитывается более 2х десятков), которые с 2014 г.

будут выходить исключительно в электронном варианте на платформе Open Journal System. Данный проект также уже начал реализовываться, центр ответственности за его реализацию – Научная библиотека. Кстати, 3 журнала КФУ сегодня входят в перечень изданий, индексируемых Scopus.

Несколько особняком от представленных выше front-end решений стоит модуль Имидж-каталога, разработанный также в Центре цифровых технологий и сервисов. Имидж-каталог представляет собой on-line версию генерального алфавитного карточного каталога Научной библиотеки, который насчитывает около 1500 тыс. карточек. Особенностью данного проекта является возможность как алфавитного, так и контекстного поиска по базе карточек, который стал возможен благодаря предварительной обработке отсканированных образов библиографических карточек и присоединению текстового слоя к каждому изображению. В ближайшее время в модуле появится возможность удаленного заказа литературы по результатам поиска. Заказ будет реализован через соответствующий сервис АБИС Руслан.

Подводя итог, можно сформулировать следующие выводы: сегодня понятие электронных ресурсов библиотеки необходимо трактовать в широком смысле.

Электронный ресурс библиотеки - это любая информация, которая формируется и существует в цифровой форме, снабжена определенным набором метаданных, участвует в поиске и может быть доставлена пользователю посредством интегрирующей программной оболочки.