

Выбор ЭБС: вопросы интеграции и сервисов

Большое количество публикаций и дискуссий в последнее время посвящено оценке электронно-библиотечных систем (ЭБС) с точки зрения формальных критериев. В основном обсуждается содержимое ЭБС. Однако, при всей важности этого критерия, нельзя упускать из вида и другие параметры, которые тоже могут оказать влияние на процесс принятия решения по выбору той или иной ЭБС.

Как показал анализ, проведённый НЭ-ИКОН¹, ни одна ЭБС по своему содержанию не удовлетворяет всем потребностям вузов. Поэтому вузу, особенно крупному, необходимо подписываться на несколько ЭБС одновременно. Так, например, Санкт-Петербургский государственный университет подписывается одновременно на ЭБС «Университетская библиотека онлайн» и издания ЭБС «Лань». Будем считать ЭБС все системы, которые так себя сами называют.

Как правило, у вуза есть своя собственная электронная библиотека. Поэтому очень важным критерием ЭБС является возможность её интеграции в существующую информационную инфраструктуру вуза. В ином случае, без интеграции, искать необходимую книгу придётся последовательно во всех выписываемых ресурсах. Какие методы интеграции возможны вообще? В качестве примера приведём скриншот раздела сайта известного агрегатора — компании «Эбско» <http://www.support.epnet.com/eit/>, посвящённого интеграции (рис. 1). На нём мы

видим множество вариантов. Попробуем вкратце объяснить, что же мы видим у «Эбско». В левом верхнем углу рассказывается о поддержке протокола SOAP (Simple Object Application Protocol). С его помощью все желающие (имеющие право доступа) могут обращаться из своих программ к базе данных «Эбско» и полученные данные использовать в своих программах. Очень важна поддержка Z39.50. С его помощью можно осуществлять распределённый поиск одновременно в различных базах. Также есть возможность выгрузки всех метаданных для интеграции их пользователем с другими имеющимися у него данными. Именно этой возможностью воспользовалась библиотека РГГУ (подробности ниже) для интеграции «Эбско» со своими данными. Пользователь может использовать виджеты, встраивать к себе на сайт поисковую строку для поиска в «Эбско» и т.п.

Какие же методы интеграции предлагают российские ЭБС? Их, к сожалению, существенно меньше. Только «КнигаФонд» и «БиблиоТех» предлагают закрытый API (Application Programming Interface, или Интерфейс программиро-

Автор



Михаил Ефремович ШВАРЦМАН
начальник отдела исследования компьютерных систем Российской государственной библиотеки

Рис. 1. Интеграция в «Эбско»



¹ <http://www.aselibrary.ru/blogs/archives/764/>

вания приложений – набор инструментов, предоставляемых программным обеспечением для использования во внешних программах для организации взаимодействия) к своим фондам и список изданий в формате MARC21, содержащихся в ЭБС. Совсем недавно возможность выгрузки записей в формате RUSMARC появилась и в «Университетской библиотеке онлайн». Это хорошо, но явно недостаточно. В силу быстрого роста фондов, а рост у всех ЭБС значительный, список изданий моментально устаревает, а закрытый API требует времени и денег для его внедрения в вузовскую инфраструктуру. Большинство ЭБС не предоставляют и этого, хотя важность задач понимают все представители ЭБС. Самое простое решение этой проблемы – это предоставление издателями (агрегаторами) списка своих изданий по протоколу OAI-PMH. В настоящее время уже разработано достаточное количество свободно распространяемого программного обеспечения (ПО) для организации OAI провайдера (сервера, который отдаёт всем желающим библиографические описания в формате DublinCore по протоколу OAI-PMH). Например, это REPOX <http://www.repoх.ist.utl.pt/>, хорошо зарекомендовавший себя в работе Европейской библиотеки, есть и другие.

Для сбора данных по OAI тоже есть свободно распространяемое ПО, например, VuFind <http://www.vufind.org>, разработанное в Университете Виланова (Пенсильвания). В библиотеке Российского государственного гуманитарного университета уже есть успешный опыт установки и использования этого ПО для интеграции всех подписываемых зарубежных ресурсов и своего электронного каталога в единую поисковую систему <http://www.vufind.lib.rsuh.ru/>. В такой системе каждый вуз сможет самостоятельно определять, из каких ЭБС ему нужно получать библиографические записи, собирать их при помощи своего OAI-сборщика и загружать в свой каталог.

При рассмотрении проблемы интегрируемости необходимо решить, какую модель создания ЭБС выберет вуз. Возможны следующие подходы:

- подписка (как описано выше, вуз подписывается на различные ресурсы и создаёт средство одновременного поиска по всем ресурсам);
- вуз создаёт у себя программно-аппаратный комплекс (платформу) для последующей загрузки в неё электронных книг.

У каждого подхода есть свои плюсы и ми-

нусы. В случае подписки у вас нет организационных проблем (и весьма существенных) с установкой и поддержкой платформы. Но, с другой стороны, у вас нет и уверенности в долговременном существовании ЭБС, на которую вы подписались. Как известно, «всё проходит», коммерческая компания может разориться, может изменить политику комплектования (как показывает опыт подписки на онлайн-журналы, в год меняется до 10% состава коллекции) или ценовую политику, у вас может не быть в какой-то период денег на подписку. В случае подписки вы потеряете доступ. В случае хранения книг на своей платформе вы будете иметь их всегда. В настоящее время автору известна только одна российская компания – «БиблиоТех» <http://www.bibliotech.ru/>, предлагающая такую платформу для вуза. Некоторые ЭБС, например «Университетская библиотека онлайн», предлагают промежуточные решения, позволяющие создать свою коллекцию, но на платформе ЭБС. Это тоже возможный подход, но нужно учесть, что при этом интеграция с иными ЭБС становится невозможной.

Также нужно понять, хотите ли вы выписывать (покупать) уже готовую коллекцию, или набирать издания в свою электронную библиотеку выборочно. С одной стороны, готовая коллекция сильно экономит время библиотекаря. Как приятно, когда кто-то о тебе уже позаботился. С другой стороны, для вашего вуза, возможно, будет предпочтительным индивидуальный подход. Ряд ЭБС, например «iBooks» (<http://www.ibooks.ru>), позволяют подбирать книги индивидуально. Следует отметить, что это очень трудоёмкий путь. Справедливо отмечает Ирина Билан в своих комментариях к блогу Андрея Фёдорова про ЭБС: «Кто должен считать, анализировать, имея квалификацию при этом? Вы представляете, какой это объём работы? Кто из преподавателей бросит всё и будет заниматься анализом контента ЭБС по своей специальности? Ему надо писать УМК. И, зачастую, преподаватели не идут даже на менее затратные (с точки зрения времени) виды сотрудничества с библиотекой вуза, в котором трудятся. Или библиотекари? Но, опять же, вопрос наличия времени – такой объём работ будет реальным, только если подсчётами будет заниматься библиотекарь, который занимается только вопросами ЭБС и ничем другим. А таких ставок у нас нет. В зависимости от масштабов вуза это может быть и не один такой библиотекарь»². Трудно с ней не согласиться. Хотя, может быть, получит

² <http://www.ideafor.info/?p=3195#comment-932>



распространение какой-нибудь промежуточный вариант...

Кроме оценки возможности интеграции различных ЭБС также нужно будет оценить и качество реализации в них отдельных функций. Можно выделить следующие показатели.

Типы предоставляемых изданий. Пока ЭБС предлагают нам, в основном, только книги. В ЭБС «Университетская библиотека онлайн» есть очень небольшое количество аудиокниг и мультимедийных материалов. Хотя отдельные ЭБС и пишут о возможности хранения в них не только книг, примеров реализации этих деклараций мне найти не удалось. Попробуем сравнить это с тем, что предлагают зарубежные издатели. Вышеупомянутая компания «Эбско» сейчас продвигает на рынок свой новый продукт — коллекцию книг NetLibrary <http://www.netlibrary.com/>. Она содержит 220 тыс. электронных книг и 16 тыс. аудиокниг, при этом 6 тыс. из них могут быть прослушаны на iPod. Понимая, что до абсолютных показателей нашим ЭБС далеко, хочу обратить внимание на процент аудиокниг.

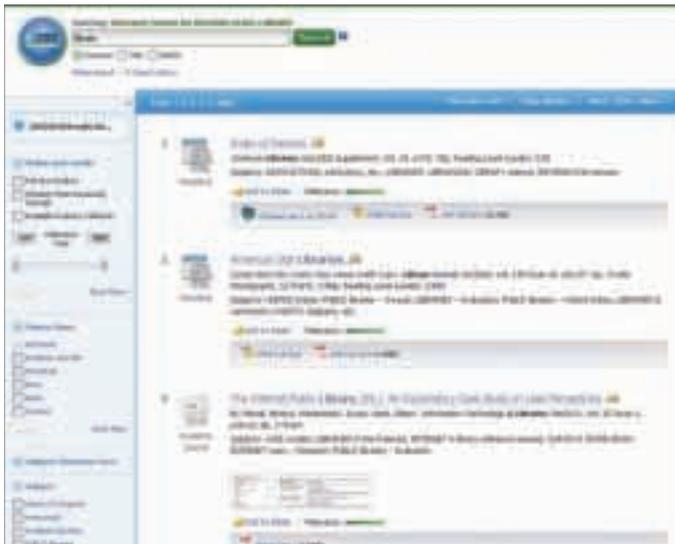
Продвинутость поиска. Во всех ЭБС поиск устроен по-разному, и возможности его очень различны. Количество поисковых полей разное (от двух полей в ЭБС «Лань» до девяти полей в IQlib), количество используемых Булевых операторов тоже разное (от одного «И» в «БиблиоТех» до трёх «И», «ИЛИ», «ИЛИ» в ЭБС book.ru). Качество морфологического анализа тоже отличается в зависимости от используемой системы полнотекстового поиска. Большое значение имеет возможность усечения справа при задании поискового термина. В некоторых системах, особенно тех, кто использует «ЯндексСервер», такая возможность отсутствует, а в некоторых, например в «Контекстум», работает также и усечение слева. Вопрос о необходимости продвинутого поиска в информационных системах постоянно обсуждается. С одной стороны, для продвинутых пользователей это очень удобно. Но, с другой стороны, нынешнее «поколение Google» привыкло искать в одной поисковой строке и не хочет осваивать «премудрости» расширенного поиска и языка запросов. Как показывают исследования, в 95% случаев читатели университетских библиотек используют для поиска в электронных ресурсах простой поиск. Поэтому прежде чем предъявлять требования по этому критерию, вы должны решить, на кого вы ориентируетесь. Полнотекстовый поиск, как правило, есть во всех ЭБС

и работает примерно одинаково (с точностью до морфологии), но в некоторых, например в «IPRbooks», предусмотрена возможность поиска не во всей ЭБС, а только в той книге, которая загружена в просмотрщик полного текста, и реализуется средствами просмотрщика. Это, конечно, не так удобно.

Используемые рубрикаторы. Обычно в каждой вузовской библиотеке есть система классификации. Чаще всего это ББК или УДК. В большинстве ЭБС используется своя собственная система классификации и поэтому проблема совместимости с системой классификации вуза становится особенно сложной. Исключением являются «IQlib», использующая ББК и УДК, и «Контекстум», использующий УДК. Для меня непонятен такой подход ЭБС, ведь практически все учебники, попадающие в ЭБС, имеют индекс ББК, пусть краткий, для массовых библиотек, но индекс. Если в ЭБС делается библиографическое описание книги, то почему бы туда не вставить этот индекс? Он позволит автоматически строить иерархию тем в коллекции. Надеюсь, что создатели ЭБС тоже прочитают эту статью (как показал анализ форумов про ЭБС, наибольшее количество комментариев там оставляют не библиотекари, а создатели ЭБС) и возьмут этот совет в работу.

Визуализация результатов поиска. При оценке методов вывода результатов следует обратить внимание на два момента: наличие сниппетов и возможность искать в найденном. Сниппеты (цитаты, содержащие поисковый термин) могут выводиться по-разному (сниппетов нет вообще, выдаётся только первый сниппет, последовательно выдаются все сниппеты). Для оценки правильности работы системы и верности составления запроса сниппеты просто необходимы. Для обеспечения возможности уточнения поискового запроса чаще всего используется кластеризация (разбиение результатов на классы) выдачи, практически нигде не реализованная в российских ЭБС. В процессе поиска должно проходить взаимодействие пользователя и системы. Получив много «шума» в ответ на первый вариант запроса, пользователь должен получить от системы общую картину результатов, чтобы понять, в каком направлении он может уточнить запрос. Для этого система должна разбить список полученных документов на классы по различным признакам (год издания, издатель, тема и т.п.) с указанием найденных в каждом классе результатов. Такой фасетный поиск (facet search) давно уже используется

Рис. 2. Фасетный поиск в EDS



во многих зарубежных поисковых системах, например в упомянутой выше VuFind, или в Solr, или на платформах некоторых издателей, например EBSCO Discovery Service (EDS) и т.п. На рис. 2 изображён скриншот, на котором видны возможности предоставляемые EDS.

Читатели, долго занимающиеся информационными системами, помнят, что этот термин в нашей стране был известен очень давно, но почему-то в последние годы практически не использовался.

Международный консорциум EIFL организовал продвижение свободного программного обеспечения с открытым кодом для организации фасетного поиска (Blacklight, VuFind and Apache Solr). Для этого был запущен проект Faset search <http://www.eifl.net/faceted-search>, в рамках которого проводились международные онлайн-семинары. В некоторых автор участвовал и может засвидетельствовать их полезность для освоения этих продуктов. Хотя, опять-таки, эта информация нужна больше для создателей ЭБС.

Можно сформулировать ещё множество критериев оценки. Например, качество библиографического описания электронных изданий. Хотя, по моим наблюдениям, оно во всех ЭБС не очень хорошее и выбирать тут особенно не приходится. Наличие постоянного адреса (URL) документа для ссылки на него тоже очень важно для интеграции ЭБС. Большинство ЭБС всем документам присваивают постоянный адрес, на который можно сделать ссылку в статье и быть уверенным, что она не поменяется, по крайней мере, пока существует эта ЭБС.

Можно сравнить методы вывода полного текста: насколько для ваших читателей удобно работать во встроенном флэш-плеере или для них предпочтительнее отдельно приобретаемый просмотрщик полных текстов типа Vivaldi, который, с одной стороны, требует дополнительных затрат, но, с другой стороны, обеспечивает дополнительную функциональность.

Большое значение имеют права по копированию, которые по-разному предоставляют нам ЭБС. Западные издатели тоже по-разному решают проблему копирования. Так например, Ebrary показывает книги постранично, в виде картинки, в специальном просмотрщике, а Elsevier публикует свои книги в текстовом виде в формате pdf и никак технически не ограничивает возможность копирования. Elsevier, конечно, проще — на него не распространяется 4-я часть ГК РФ, да и он, вообще, хозяин своих книг. Российским ЭБС сложнее, за ними законодательно закрепили обязанность исключить возможность копирования. Но не запретили цитирование, а его все понимают по-разному. Некоторые вообще не разрешают ничего копировать даже для цитирования, в некоторых мы можем отобрать до 10% страниц из книги, в некоторых до 1000 знаков. Поскольку при составлении научной работы переписывать ручкой с экрана — не самое приятное занятие, обращайтесь внимание при выборе ЭБС и на эту деталь.

Для получения общей статистики библиотеки вуза библиотекарю приходится сводить воедино статистику использования от каждого выписываемого ресурса. Удобнее всего, если они будут в одинаковом формате. Ряд ЭБС выдают некоторую статистику, но у каждой свой набор таблиц и для сведения их в единую форму библиотекарю приходится затратить много ручного труда. Опять обращаясь к зарубежному опыту, видим, что уже давно существует стандарт предоставления библиотечной статистики COUNTER <http://www.projectcounter.org>, соблюдаемый многими зарубежными издателями. К сожалению, из всех российских издательств только компания «ИВИС» следует этому стандарту. Поэтому одним из критериев оценки ЭБС должна стать оценка удобства их статистики.

Список критериев этим не исчерпывается, и автор уверен, что у каждого вуза он будет свой со своими весами. Задачей этой статьи не было посоветовать какую-то конкретную систему, а лишь обратить внимание библиотекарей на важные моменты при выборе ЭБС. ■