



World Library and Information Congress: 69th IFLA General Conference and Council

1-9 August 2003, Berlin

Code Number: 155-R
Meeting: 86. Bibliography
Simultaneous Interpretation: Yes

Будущее (электронных) национальных библиографий

Юха Хакала

Библиотека Университета Хельсинки
Национальная Библиотека Финляндии

(Juha Hakala)

Резюме

Возрастает число национальных библиографий доступных в Интернет и отражающих содержание сети Интернет. В представленном докладе анализируется воздействие произошедших (или возможных) изменений на функции, которые выполняют базы данных национальных библиографий. Для копирования каталогизационных записей, доступных в Интернет, необходим протокол Z39.50. Каталогизация сетевых ресурсов проанализирована с точки зрения традиционных библиотечных систем, информационных поисковых порталов и систем управления цифровыми объектами.

Введение

В январе 2004 г. Финляндия станет третьей страной в мире, в которой национальная библиография будет доступна только через Интернет. Издание национальной библиографии в печатном виде было приостановлено еще несколько лет назад, а в конце 2003 г. и CD-ROM-версия базы данных нашей Fennica станет историей. CD-ROM был необходим, в первую очередь, для копирования каталогизационных записей, с тех пор, как клиенты внутри страны, так и за рубежом, получили возможность копировать в режиме он-лайн через протокол Z39.50, то необходимость поддерживать CD-ROM-версию отпала. Многие страны в скором времени окажутся в том же положении; издание национальной библиографии на любом материальном носителе является только данью традиции.

В первую очередь в докладе будут проанализированы предпосылки прекращения публикации национальной библиографии на любом материальном носителе. Затем, используя в качестве

отправной точки финский опыт и планы, будут обсуждены пути развития национальных библиографий с учетом увеличивающегося объема каталогизации электронных ресурсов. В этом контексте важное значение на пути создания национальной библиографии и услуг, оказываемых на ее основе, будет иметь реализация плана создания финского национального информационно-поискового портала и системы управления университетских библиотек цифровыми объектами.

Доступ к базам данных национальной библиографии

Согласно обзору, проведенном Унни Кнутсен в 2001 г., через Интернет были доступны 23 национальные библиографии. Кроме того, 15 национальных библиографических агентств планировали к 2006 г. предоставить такой доступ; из них, по крайней мере, 9 уже выполнили эту задачу (Knutzen). Число национальных библиографий, доступных через Интернет, растет очень быстро, сетевой доступ становится преобладающим средством доступа.

Это, конечно, очень положительная тенденция, но этого недостаточно. Осуществление поиска в базах данных национальных библиографий только через сеть не может удовлетворить библиографов, т.к. копирование каталогизационных записей будет невозможно, если не будет поддерживаться информационный поисковый протокол Z39.50. Все усилия должны быть направлены на совершенствование каталогизации. Каталогизация только в том случае будет эффективна, если полученными результатами смогут пользоваться и другие; для этого необходимо выполнять технические и политические соглашения.

При проведении следующего анализа современного состояния электронных национальных библиографий надо будет учесть наличие возможности копирования каталогизационных записей. Это можно сделать, если добавить несколько вопросов, адресованных национальным библиографическим агентствам, обеспечивающим доступ через Интернет:

1. Поддерживается ли протокол Z39.50?
2. Если да, то в каком формате(ах) и каком виде могут передаваться данные?
3. Является ли доступ к Z39.50 свободным или необходимо заключить договор? Если да, то какая оплата?
4. Хотели ли бы Вы договориться о взаимном обмене данными с другими национальными библиографическими агентствами на бесплатной основе?
5. Если доступ Z39.50 отсутствует, то доступны ли ваши национальные библиографические данные или часть их через другие системы с целью копирования каталогизационных записей?
6. Если да, то каковы условия пользования этой базой данных?

Национальные библиографические агентства, обеспечивающие доступ через Интернет, могут и должны предоставлять в сетевом режиме через протокол Z39.50 свои библиографические записи в формате MARC с целью копирования каталогизационных записей. Кроме того имеются достаточно полные сводные каталоги, например WorldCat (OCLC), насчитывающий к концу 2002 г. около 50 миллионов записей, и в котором приблизительно 60.000 записей содержали сведения об изданиях опубликованных в Финляндии. Но наши национальные библиографии обычно имеют более полный охват изданий, чем международные сводные каталоги, и качество записей в них не так высоко. Тот, кто заинтересован в более полном охвате финской литературы или в высоком качестве записей, должен использовать финскую национальную библиографию. Эта ситуация может измениться, если записи Fennica будут загружены, например, в WorldCat. Некоторые национальные библиографические агентства уже сейчас выбрали этот путь развития. Интересно было бы узнать, есть ли у НБА планы по более тесному сотрудничеству, например, с OCLC.

Z39.50 и национальные библиографии

В настоящее время поддержка Z39.50 является стандартной функцией в каждой хорошо интегрированной библиотечной системе и IR портале типа MetaLib. Проблема заключается в том, что, несмотря на существующие одобренные международным сообществом рекомендации по использованию Z39.50 – Bath profile – интегрированные библиотечные системы до сих пор не могут полноценно «разговаривать» друг с другом. Для этого имеется много причин, включая и следующие:

- Не все национальные библиографии используют MARC21 (формат, находящийся под мандатом Bath profile). Некоторые из тех, кто не использует его, не могут конвертировать записи в MARC21, функционирующим на основ Z39.50, специально предназначенном для обмена данными. Так как инструментарий по преобразованию формата не содержит функции копирования каталогизационных записей, то это является серьезной проблемой. Лучшим средством решения этой проблемы является использование хорошего конвертера формата MARC типа USEMARCON (USEMARCON) в качестве составной части процесса копирования каталогизационных записей. Тогда, даже если скопированные записи доставляются в “неправильном” формате MARC, то их легко можно будет конвертировать в местный формат MARC. Все, в чем Вы нуждаетесь, это набор таблиц преобразования, которые при наличии эксперт формата USEMARCON, который знаком с обоими форматами – исходным и целевым – может настроить их в течение нескольких недель.
- В форматах MARC может быть использован расширенный набор символов, и записи в формате MARC21, например, кодируются различными способами. В конечном счете, UNICODE сможет изменить эту ситуацию, но пройдет еще немало времени прежде, чем каждая библиотечная система начнет поддерживать UNICODE в полном объеме. Z39.50 имеет блестящее решение этой проблемы: обсуждение набора символов. Но только несколько библиотечных систем поддерживают эту функцию в службе инициализации в Z39.50. Быстро решить эту проблему можно при использовании USEMARCON или другого приложения, способного преобразовать соответствующим образом набор символов.
- В зависимости от целей Z39.50 обеспечивает различную семантику (поисковые образы). Bath profile дает четкие рекомендации по поддержанию определенных стандартом способов выражения запроса и терминологических сочетаний в контексте Z39.50. Национальным библиотекам нет никакого смысла не придерживаться рекомендаций данного профиля; исключение может быть только в том случае, когда в базах данных национальной библиографии приводится большое количество элементов и тогда было бы лучше использовать дополнительные точки доступа.
- В зависимости от целей и состава клиентов Z39.50 поддерживает различные наборы услуг и функций, определенные в стандарте. Лучший способ избежать неприятностей в данной области – придерживаться Bath profile. К счастью, для создания копий каталогизационных записей не требуется экзотических средств; для этого достаточно таких базовых функций как инициализация, поиск и представление.

В качестве примера можно рассмотреть ситуацию в Финляндии. Наша база данных национальной

библиографии базируется на MARC21-FIN, который является финской расширенной и усовершенствованной версией MARC21. Университетские и политехнические библиотеки используют тот же самый формат, а публичные библиотеки все еще работают на основе старого формата FINMARC и набора символов 6937/2 ИСО. Однако, публичные библиотеки не испытывают проблем при копировании каталогизационных записей из Fennica.

Поставщики крупнейшей финской библиотечной системы интегрировали USEMARCON в каталогизационную систему клиентов и внесли некоторые усовершенствования в приложение. При использовании публичной библиотекой любой из этих систем копирование записей из национальной библиографии в формат MARC21-FIN через Z39.50 происходит автоматически: USEMARCON конвертирует эти записи в FINMARC и 6937/2 ИСО. В результате обращение к сетевому доступу выросло, а подписка на CD-ROM Fennica – уменьшилась. Поэтому было принято решение прекратить выпускать Fennica на CD-ROM; в нашей случае теперь нужна только сетевая версия национальной библиографии. Другим феноменальным фактором стало быстрое совершенствование компьютерной технологии; 10 лет назад не любая публичная библиотека могла себе позволить купить сервер для копирования каталогизационных записей; теперь это вообще не является проблемой.

С 2004 г. иностранные пользователи одновременно будут получать MARC21 с Z39.50. USEMARCON также внедрен в сервер Voyager Z39.50, и теоретически это позволяет нам распространять библиографические записи в любом формате MARC и с любым набором символов. Реально только небольшая часть конверсионных форматов может быть поддержана любым НБА, поэтому для поддержания таблиц преобразования необходимо детально изучить как форматы создателя, так и получателя. Используя MARC21 в качестве общего для НБА формата можно было бы создать большое разнообразие способов конвертирования, и тогда любая национальная библиотека через Z39.50 могла бы напрямую копировать записи в формате MARC из любой национальной библиографии.

Использование Z39.50, снабженного программным продуктом типа USEMARCON для преобразования формата и набора символов, позволяет осуществлять эффективный информационный поиск и копирование каталогизационных записей из национальных библиографий. В связи с тем, что технология развивается очень быстрыми темпами, возникла необходимость пересмотреть рекомендации по национальным библиографическим службам и включить в них руководства по Z39.50. Удивительно, что этот стандарт, или его ИСО версия, ИСО 23950, даже не упомянуты в Заключительных рекомендациях Международной Конференции по Национальным Библиографическим Службам (2002 г.).

Для поддержания технической инфраструктуры между НБА необходимо заключать соглашения по обмену данными. Международное сотрудничество типа Скандинавского виртуального сводного каталога (<http://www.lib.helsinki.fi/svuc/>), объединяющее Скандинавские национальные сводные каталоги в единую виртуальную службу, или Ассоциация ONE (<http://www.oneassociation.org>), ставят своей целью сделать библиографические данные, в том числе и записи национальной библиографии, включенные в сводный каталог, доступными для международного сообщества.

Обычно национальные библиографические агентства редко используют библиографические записи, созданные другими агентствами, поэтому надо расширять совместное использование этих данных. Я думаю, Ассоциация ONE является подходящей моделью для сотрудничества по использованию записей. С другой стороны, совместное использование НБА авторитетных данных является также не плохой идеей, надо стараться создать так называемые Виртуальные

международные авторитетные файлы. Но хотя версия 2.0 Bath profile предоставляет возможность найти с помощью Z39.50 авторитетные данные, библиотечные системы еще не имеют этого приложения к Z39.50, которое поддерживало бы эту функцию, не говоря уже о других отсутствующих технических и политических усовершенствованиях.

Пример обмена авторитетными записями показывает, что национальные библиотеки, похоже, имеют, соответствующий специальный Z39.50 и другие средства, но другие библиотеки, как правило, не имеют их. Надо формировать соответствующие группы библиотек, чтобы развивать те функции, которые наиболее важны для обмена данными. По крайней мере, одна подобная группа уже есть National Libraries' Endeavor Advisory Board, NLEAB (<http://www.infotoday.com/IT/nov01/news18.htm>).

Что такое протокол ZING?

Многие из Вас, возможно, слышали о протоколе ZING (ZING). Этот акроним означает Z39.50 следующего международного поколения. Технически он существенно отличается от Z39.50, хотя с точки зрения пользователя выглядит точно также. Поддерживает этот протокол Библиотека Конгресса. Эти два протокола мирно сосуществуют между собой.

Проще говоря, протокол ZING это Z39.50 для Интернет. Используя, например, услугу ZING SRU можно посылать запросы, кодируемые в URL, подобно такому:

```
http://www.loc.gov/sru.cgi?query=dc.title=cat&maximumRecords=10&recordSchema=DC&sortKeys=title,dc&startRecord=1
```

Эта технология работает, несмотря на ужасный внешний вид приведенных данных. Например, Google использует подобные технические решения, и также в OpenURL библиографические данные о ресурсе посылались в распознавательную систему OpenURL, находящую в URL. ZING сервер назад посылает записи не в обменный формат ИСО 2709, а кодирует в XML, используя схему Дублинского Ядра.

Пока заменять в библиотечных системах протокол Z39.50 на ZING не стоит. Но все-таки полезно поддержать оба протокола Z39.50 и ZING, чтобы сделать библиотечные системы более доступными для внешнего пользователя. Например, если бы база данных национальной библиографии была доступна через ZING, издателям и другим организациям, осуществляющим мемориальные функции (музеи, архивы), было бы легче подключить свои системы к национальной библиографии.

По моему мнению, НБА должны способствовать распространению национальных библиографических данных путем создания наших записей, которые будут доступны в XML синтаксисе через ZING и возможно также через Открытую Инициативу Архивов (<http://www.openarchives.org/>). Через некоторое время наши библиотечные системы будут иметь эти функциональные возможности; и к тому времени мы должны быть готовы использовать эти технологии.

Будущее национальных библиографий

Согласно обзору Унни Кнутсен, представленному ранее, национальные библиографии могут (или нет) слиться с ОПАС (сетевой публичный доступ к каталогу) национальных библиотек или

другими библиографическими системами, но всегда считалось, что национальная библиография базируется на объединенной библиотечной системе и с технической точки зрения представляет собой единую базу данных.

Поставщики библиотечной системы и другие компании развивают информационные поисковые порталы (IR), и наиболее известным примером является MetaLib, и системы управления цифровыми объектами или DOMS, такие как ENCompass от Endeavor Information Systems.

IR портал может быть рассмотрен как расширение традиционного OPAC. Обычно портал позволяет вести одновременный поиск в нескольких целевых базах данных, используя различные протоколы запроса, синтаксическое написание и наборы символов. Это позволяет создателям портала хранить запросы и просить систему периодически их выводить. В последнее время стало нормой, что порталы поддерживают динамическое соединение с OpenURL. Хотя эти услуги могли бы быть интегрированы в традиционную библиотечную систему, поставщики обычно выводили их в отдельные приложения. Это зависит от некоторых технических и большого количества коммерческих причин. Библиотека, использующая ILS от поставщика А может покупать портал от поставщика Б, если ей не надо менять библиотечную систему.

Те же самые технические и коммерческие принципы действуют и применительно к DOMS; библиотечная система могла бы содержать требуемые функциональные возможности. Но с тех пор как последние разработки DOMS имеют дело с разнообразными документами и форматами метаданных и контролем над защитой предоставляемой информации, традиционная библиотечная система должна быть сильно расширена, чтобы соответствовать всем требованиям.

Когда библиотеки организуют портал и DOMS системы, одна интегрированная система может быть заменена двумя или тремя приложениями, связанными посредством стандартных интерфейсов. В финском сети Linnea в 2004 г. мы будем в качестве библиотечной системы использовать Voyager, MetaLib как IR портал и ENCompass как DOMS.

С точки зрения каталогизации, появление этой новой технической среды интересно по двум причинам. Во-первых, ссылки на информацию, которая представляет собой описание ценных электронных ресурсов, должны быть помещены не в традиционной библиотечной системе, а на портале. Во-вторых, для эффективного использования портала и DOMS потребуются новый вид метаданных, которые пока не обеспечены ни новыми правилами каталогизации, ни форматами. Эти изменения повлияют и на OPACs и на национальные библиографии.

Позвольте мне более подробно остановиться на первой проблеме. Для обеспечения связи между системой, основанной на MARC21, и сетевыми ресурсами надо поместить URL сетевых ресурсов в 856 поле записи MARC. Каждый раз при изменении URL мы должны вносить изменения в записи в каждой базе данных, из которых запись может быть скопирована.

К сожалению, по крайней мере, в Финляндии только некоторые библиографические базы данных содержат URL, которые большей частью не работают, или в худшем случае отсылают от записи MARC к совершенно другому источнику.

Решение, основанное на поле 856, также не способно решить проблему копирования; то есть это означает, что может быть много копий одного и того же источника, размещенного в разных местах, и один пользователь может обратиться к одной копии, другой пользователь – ко второй копии, третий – к третьей. Практически такой ситуации быть не может, т.к. в дальнейшем доступ будет осуществляться на основе лицензионных соглашений.

Подобная проблема копирования может быть решена с помощью поддержания упрощенных URL, на основе OpenURL. В данном случае, ссылка и информация о лицензии хранится в базе данных распознавателей OpenURL; в Финляндии это находится в системе SFX, которая включена в портал MetaLib. После того, как это сделано, мы должны удалить данные отсылки из национальной библиографии. Нет смысла поддерживать статичные ссылки национальной библиографии и динамичные ссылки в БД распознавателей OpenURL.

Наш вывод заключается в том, данные национальной библиографии должны быть поделены между традиционной библиотечной системой и системой типа SFX OpenURL. Некоторые метаданные, имеющие отношение, например, к долгосрочному хранению могли бы располагаться в системе управления цифровым объектом. Отсюда логически вытекает следующий вопрос, какие метаданные должны относиться и к БД распознавателей OpenURL и к национальной библиографии?

Возможно, большинство каталогизаторов согласятся, что перемещенные или скопированные из библиотечной системы в распознавательную систему OpenURL элементы данных, продолжают оставаться частью национальной библиографии. Содержание поля 856 - пример этому. Вызывают проблемы только те элементы, которые не были использованы. Должна ли в национальной библиографии содержаться только информация о лицензированных электронных ресурсах? Я считаю, что, безусловно, «да»; когда наши системы развиваются таким способом, что в них можно хранить и использовать лицензированную информацию, то мы должны, по крайней мере, использовать эти данные. Например, время и условия использования депонированных ресурсов должны быть указаны в национальной библиографии.

Так, в ближайшем будущем наши национальные библиографии смогут состоять из традиционной библиотечной системы и распознавательной OpenURL. С точки зрения конечных пользователей эти два приложения, объединенные в единую услугу, обеспечивают динамичную связь с электронным содержанием. Это содержание может храниться в системе управления цифровыми объектами, которая явится третьим краеугольным камнем наших будущих служб. Мы еще полностью не понимаем всего назначения новой прикладной среды; например, мы не знаем, как поставлять библиографические записи вместе с ссылкой на информацию, если эти данные находятся в двух приложениях. Эти проблемы должны обсуждаться и, в конечном счете, решены, например, на ИФЛА.

References

Bath profile home. The National Library of Canada [on line]. <http://www.nlc-bnc.ca/bath/> (2003-07-10)

Knutsen, Unni: Electronic national bibliographies: state of the art review [on line]. <http://www.ifla.org/IV/ifla69/papers/109e-Knutsen.pdf> (2003-07-10)

OCLC WorldCat home [on line]. OCLC Online Computer Library Center. <http://www.oclc.org/worldcat/> (2003-07-10)

USEMARCON home [on line]. The British Library. <http://www.bl.uk/services/bibliographic/USEMARCON.html> (2003-07-10)

Z39.50 Maintenance agency home page. Library of Congress [on line]. <http://lcweb.loc.gov/z3950/agency/> (2003-07-10)

ZING home page. Library of Congress [on line]. <http://www.loc.gov/z3950/agency/zing/zinghome.html> (2003-07-10)