



68th IFLA Council and General Conference

August 18-24, 2002

Code Number:	023-144-R
Division Number:	IV
Professional Group:	Classification and Indexing with Cataloguing - Workshop
Joint Meeting with:	CILIP Cataloguing and Indexing Group
Meeting Number:	144
Simultaneous Interpretation:	-

Общий Интернет-портал для библиотек, архивов и музеев – BAM-портал (<http://www.bam-bw.de>)

Геральд Майер
Управление архивами Земли Баден-Вюртемберг
Штуттгарт, Германия

Введение

Обязанность архивов, библиотек и музеев перед обществом в соответствии с законодательными актами состоит не только в том, чтобы сохранить культурное достояние, но и сделать его доступным обществу. Для доступа и предоставления информации возрастающе важную роль играют как онлайн-средства, так и Интернет. Три типа организаций – архивы, библиотеки, музеи – развивались в прошлом по различным принципам и поэтому применяли различные методы раскрытия содержания документов и их предоставления. В современном информационном обществе пользователи, а еще более учреждения, оказывающие информационные услуги, нуждаются в расширении возможностей предметно-тематического поиска, поэтому создание сети содержательно взаимосвязанных источников из информационных учреждений трех типов отвечало бы новым требованиям. Неудивительно поэтому, что в последнее время возникают инициативы, программы и проекты, занимающиеся проблемой виртуального объединения фондов данных информационных учреждений трех типов. Особенно форсируется в рамках плана акций Европейского Союза "eEurope" разработка сети с участием многих европейских стран для преобразования в цифровую форму политематического культурного наследия. На национальном уровне тоже делаются попытки в этом направлении. Так, Управление архивами Земли Баден-Вюртемберг, Баден-Вюртембергский центр библиотечного обслуживания в Констанце и Музей техники и труда в Мангейме с февраля 2001 г. разрабатывают под патронажем Немецкого исследовательского общества проект под названием "Общий Интернет-портал для библиотек, архивов и музеев. Онлайн-информационная система", кратко BAM-портал, который будет здесь представлен подробнее.

Целью этого проекта, который после года работы в конце апреля 2002 г. будет пролонгирован, является развитие внедряемых в качестве модели технических методов, которые дадут возможность упомянутым организациям в общем онлайн-портале осуществлять поиск их цифровых ресурсов (библиографической информации; данных, раскрывающих содержание; цифровых объектов). При этом развивается формат метаданных, а также предоставляется в цифровой форме культурное достояние по выбранной тематике из ретроспективных архивных, библиотечных и музейных фондов.

Исходное состояние

В целом Интернет представляет собой значительно развившуюся коммуникационную и информационную среду. Однако поиск в Интернете вследствие отсутствия некоторых структур все еще сложен. С помощью известных глобальных полнотекстовых поисковых машин, как например Google или Altavista, целенаправленный доступ к информации часто из-за очень больших объемов "попаданий" едва ли возможен /1/. Решение этой проблемы предоставляют так называемые порталы, которые образуют единый пункт входа для определенных тематических поисковых задач. Применение одновходовой Интернет-поисковой машины для виртуального поиска только одних архивов дает недостаточные результаты. Но в отличие от архивов, а также музеев, библиотеки уже давно представлены в Интернете, в частности, в каталогах OPAC /2/ и виртуальных каталогах библиотечных союзов /3/. Причины этого различны:

1. Библиотечные фонды обычно не нуждаются в ориентированном на происхождение раскрытии содержания с необходимостью тектонической связи. Следовательно, для поиска годятся решения баз данных без необходимости контекстноориентированной навигации.
2. Для раскрытия содержания и обмена данными существуют стандарты (например, MAB2, Dublin Core, Z39.50) /4/.

В противоположность этому мир архивов и музеев очень однороден. Но едва ли существуют признаваемые всеми стандарты унифицированного раскрытия содержания и форматы обмена данными – слишком различны традиции, особенно традиции управления отдельными архивами и архивными ландшафтами. В американском пространстве в качестве стандарта имеется "Кодируемое описание архивов" [Encoded Archival Description (EAD)] /5/. EAD был признан стандартом в 1998 г. и используется в США и частично в англоязычных странах для раскрытия содержания книг при онлайн-поиске. В немецком архивном деле не существует единого стандарта для раскрытия содержания документов и представления результатов в онлайн. Однако требуется, чтобы архивы также, как и другие учреждения, предоставляющие информационные услуги в современном информационном обществе, занялись проблемой представления содержания своих фондов в онлайн-среде и стремились при этом к их совместимости.

Концепция и информационные области ВАРМ-портала

Основными целями ВАРМ-портала предусматривается объединение уже имеющихся цифровых фондов организаций-партнеров – Баден-Вюртембергского центра библиотечного обслуживания, Управления архивами Земли Баден-Вюртемберг и Музея техники и труда в Мангейме – и сведения метаинформации о накопленных там данных в центральную структурируемую базу, которая образует основу для однородного поиска данных.

Пользователь Интернета сможет входить в эту базу данных с помощью Web-браузера через портал и получать содержащиеся там метаданные. Сами данные о содержании документа и цифровые

объекты, которые описаны с помощью метаданных, остаются в Интернете и тем самым на Web-сервере действующих организаций. Таким образом пользователь может одновременно получить информацию о книгах по теме, которая его интересует, о соответствующих архивах из задействованных архивов, а также информацию об экспонатах участвующих музеев.

ВАМ-портал предоставляет следующие информационные услуги:

1. Основную информацию об учреждениях-участниках.
2. Метаданные с указанием онлайн-средств поиска и онлайн-каталогов.
3. Онлайн-средства поиска и данные онлайн-каталогов.
4. Цифровые объекты.

В рамках проекта был выбран так называемый тест-контент по теме "Индустриализация" на территории теперешней Земли ФРГ Баден-Вюртемберг за период с 19 в. до начала 20 в. Сценарий поиска мог бы выглядеть, например, следующим образом: пользователь интересуется темой "Eisenbahn in Baden-Württemberg" ["Железные дороги в Баден-Вюртемберге"]. Тогда пользователь должен провести минимум три поиска с помощью самостоятельного поискового инструмента, чтобы найти все материалы по теме, которые предлагаются в различных организациях:

- Поиском в Сводном каталоге Западногерманского библиотечного союза он найдет среди прочих издание Raimund Kolb "*Baehnle*", "*Muehle*", *Zug und Bus: die Bahn im mittleren Schussental*.
- Поиском в фонде Государственного архива в Зигмарингене он найдет искомые цифровые записи соответствующих строительных актов: *Erbauung einer Nebeneisenbahn von Schussenried nach Buchau 1880, 1896-1897, Darin: Pläne*.
- Наконец, в фонде Музея техники и труда он найдет: *Dampflokomotive Eschenau T3 Cn2t*, который применялся в этот период времени, возможно, даже на этом участке дороги.

ВАМ-портал позволит провести пользователю только один одновременный поиск по ключевому слову "Eisenbahn", чтобы получить в общем перечне результаты поиска по всем трем объектам.

Раскрытие содержания, форматы данных и обмен данными

Аналогично различным задачам и материалам библиотек, архивов и музеев существуют различные методы раскрытия содержания. Это ведет к тому, что информационные системы и онлайн-представление информации у отдельных партнеров по проекту также различны:

1) Для Центра библиотечного обслуживания основой является Сводный каталог Юго-западного немецкого библиотечного союза, который доступен в онлайн-режиме через OPAC /6/. Раскрытие содержания осуществляется по библиотечным правилам, по Правилам алфавитной каталогизации в научных библиотеках [Regelwerk für die alphabetische Katalogisierung in wissenschaftlichen Bibliotheken (RAK-WB)] /7/. Для представления данных применяется Библиотечный стандарт для обмена машиночитаемыми данными [Maschinenlesbare Datenaustauschformat für Bibliotheken – MAB].

2) Раскрытие содержания архивов производится структурированно по сведениям о происхождении [т.е. сведениям об источнике, откуда поступили на хранение архивные материалы]. В Государственном управлении архивами Земли Баден-Вюртемберг оно осуществляется в виде описания архивов под заглавием с ориентировкой на базу данных, применяющую программное обеспечение MIDOSA /8/. Генерирование онлайн-средств поиска и представление данных осуществляется с применением MIDOSA в режиме онлайн и специальных разработанных Управлением архивами Земли Баден-Вюртемберг модулей представления архивных материалов /9/.

3) Музей техники и труда имеет в своем распоряжении базу данных о своих фондах с программным обеспечением FAUST /10/, которая, правда, в настоящий момент недоступна в сети. Содержание документов в музее раскрывается специалистами-документалистами по документальным правилам. При этом применяются нормативные данные из Базы данных нормативных предметных рубрик (SWD).

Существенное преимущество проекта заключается в том, что эти информационные системы работают с различными форматами данных, различными методами описания и, соответственно, раскрытия содержания, которые, по мере надобности, в зависимости от значимости соответствуют имеющемуся материалу. Но из-за предметного и структурного различия эти форматы не могут прямо служить основой для общего поиска.

Целью ВAM-проекта является не только разработка общих принципов и, соответственно, общего формата для раскрытия содержания, который сможет удовлетворить все требования описания объекта в любой отрасли и полностью заменить различные форматы, возникшие вследствие теперешних отраслевых традиций. Должен быть найден также общий знаменатель, который сделает возможным комфортабельный поиск в общем портале. Это означает, что с помощью общего формата обмена предполагается создать общую базу, содержащую ядро описания данных, отражающих фонды всех участников проекта. В качестве формата обмена данными применяется формат метаданных "DLmeta"/11/, который, со своей стороны, основывается на распространенном формате Дублинское ядро /12/.

Предпосылкой для осуществления осмысленного поиска в рамках общего ВAM-портала, выходящего за рамки отдельного фонда, является, прежде всего, осмысленное отражение применяемых форматов MAB, MIDOSA и FAUST в формате DLmeta и лежащем в его основе Дублинском ядре. При этом менее проблематичным является техническое конвертирование отдельных баз данных, чем разработка таблиц соответствия семантически адекватных полей отдельных форматов формату DLmeta. Записи в форматах MAB, MIDOSA и FAUST конвертируются в формат DLmeta-XML, затем передаются и загружаются в центральную базу данных.

Наряду с конвертированием полей из различных форматов данных определенный барьер представляет применение в организациях-участницах проекта различных тезаурусов и нормативных данных. Как Центр библиотечного обслуживания, так и Музей техники и труда используют Базу данных нормативных предметных рубрик (SWD) и Базу данных имен лиц Немецкой библиотеки (PND) /13/. Так как раскрытие содержания архивов не основывается на вербальном индексировании, то и сопоставимый перечень здесь не находит применения. Раскрытие содержания и осуществление поиска производится обычно в архивах на основе так называемого принципа происхождения. В рамках проекта апробируется применение нормативных данных и дополнительное ручное и автоматическое вербальное индексирование, так сказать, как дополнение к традиционному принципу раскрытия содержания архивов, чтобы дать возможность осуществления поиска в определенных тематических массивах, перешагивающих за пределы массивов одной организации.

Стратегии и функциональность поиска

ВAM-портал будет предлагать две формы поисковых функций:

1. При одной форме для осуществления полнотекстового поиска и поиска по определенным полям данных уже устанавливается поисковая машина с инклюзивной поддержкой перечнями ключевых слов и предметных рубрик.

2. При другой форме будет возможен предметно-тематический поиск в определенных тематических массивах с помощью иерархических предметных рубрик (перечней гиперсвязей), которые либо автоматически генерируются, либо создаются ручным способом.

Выводы и перспективы

На первом этапе проекта "Общий портал для библиотек, архивов и музеев. Онлайн-информационная система" разрабатывается апробируемая на модели концепция представления цифровых фондов литературы, архивных материалов и музейных экспонатов и возможности их поиска с помощью единых механизмов в Интернет-портале. Чтобы показать преимущества такого объединенного поиска, проект в его теперешней фазе ограничивается ресурсами патентов по теме "Энергия воды". По этим параметрам у всех трех партнеров по проекту – Центра библиотечного обслуживания в Баден-Вюртемберге, Управления архивами Земли Баден-Вюртемберг и Музея техники и труда в Мангейме – существуют достаточные цифровые массивы, так что на поисковые запросы находят ответы во всех перечисленных фондах.

Немецкое исследовательское общество дало согласие на продление договора на следующий год. На втором этапе проекта должны быть поставлены следующие две основные цели:

1. Расширение перечня отражаемых источников в ВАРМ-указателях и расширение/оптимизация поисковой системы.
2. Дифференцирование использования и расширение круга участников ВАРМ-портала.

Интернет-ресурсы

Сайт проекта: <http://www.bam-bw.de>

Партнеры по проекту:

- Центр библиотечного обслуживания в Баден-Вюртемберге: <http://www.bsz-bw.de>
- Управление архивами Земли Баден-Вюртемберг: <http://www.lad-bw.de>
- Музей техники и труда, Мангейм: <http://www.landmuseum-mannheim.de>

Дублинское ядро и DLMeta: <http://dublincore.org>; <http://www.dlmeta.de>

Ссылки

1. URLs: <http://www.altavista.com> и <http://www.google.com>.
2. ОРАС=Online Public Access Catalogue.
3. Примеры виртуальных библиотечных союзов в Интернете: Юго-западный немецкий библиотечный союз (URL: <http://www.bsz-bw.de/cgi-bin/opacform.cgi>) или Карлсруйский виртуальный каталог (KVK) (URL: <http://www.ubka.uni-karlsruhe.de/kvk.html>).
4. В ANSI/NISO Z39.50 речь идет о стандартизованном базирующемся на системе "клиент-сервер" протоколе для коммуникации между базами данных библиотечных систем (сервером) и программами доступа (клиентами) (см. <http://lcweb.loc.gov/z3950/agency/>).
5. См. также <http://www.loc.gov/ead>.
6. См. URL: <http://www.bsz-bw.de/cgi-bin/opacform.cgi>.
7. Введение и основы: <http://www.christian-kirsch.de/rak/index.html>.
8. MIDOSA см. <http://www.midosade.de>.
9. Модули были разработаны в рамках проекта "Workflow und Werkzeuge für die Bereitstellung grösserer Mengen von Archivgut" (URL: <http://www.lad-bw.de/workflow>). Результаты проекта

отражены в следующей публикации: Gerald Maier/Thomas Fricke (Hgg.). Kulturgut aus Archiven, Bibliotheken und Museen im Internet – Neue Ansätze und Techniken, Stuttgart 2002 (Werkhefte der Staatlichen Archivverwaltung Baden-Württemberg, Serie A Landesarchivdirektion). Управление архивами Земли Баден-Вюртемберг могло получить опыт оцифровки архивных материалов уже в предыдущем проекте "Цифровые формы конвертирования" (URL: <http://www.lad-bw.de/digpro>). Результаты этого проекта подробно представлены в статье: Hartmut Weber/Gerald Maier (Hgg.), Digitale Archive und Bibliotheken. Neue Zugangsmöglichkeiten und Nutzungsqualitäten, Stuttgart 2000 (Werkhefte der Staatlichen Archivverwaltung Baden-Württemberg, Serie A Landesarchivdirektion, Heft 15).

10. Документация по программному обеспечению FAUST см. URL: <http://www.land-software.de>.
11. DLmeta см. URL: <http://www.dlmeta.de>.
12. Дублинское ядро см. URL: <http://purl.org/de>.
13. SWD и PND см. URL: <http://www.ddb.de/professionell/swd.htm> и <http://www.ddb.de/professionell/pnd.htm>.