

НАЦИОНАЛЬНАЯ ИНФОРМАЦИОННАЯ ИНФРАСТРУКТУРА НАУКИ, КУЛЬТУРЫ И ОБРАЗОВАНИЯ: ПРЕДСТАВЛЕНИЕ РЕСУРСОВ ЭЛЕКТРОННЫХ БИБЛИОТЕК

Е.Н.Филинов, А.В.Бойченко

Институт системного программирования РАН, 109004, Москва,

Б. Коммунистическая, 25

boytchen@ispras.ru

NATIONAL INFORMATION INFRASTRUCTURE OF A SCIENCE, CULTURE AND EDUCATION: REPRESENTATION OF RESOURCES OF ELECTRONIC LIBRARIES

E.Filinov, A.Boychenko

Institute for system programming of the Russian Academy of Science, 109004,

Moscow, B.Kommunisticheskaya, 25

In the report are given questions concerned the choice of international standards for electronic libraries construction:

The alternative variants of choice standards on electronic catalogs and metacatalogs formats;

- The alternative variants of choice standards for interlibrary loan;
- The variants of choice standards for information search in online public catalogs (OPAC) and bibliographic databases;
- The variants of choice standards for file representation of full text documents, vector and raster images, audio- and video-materials, text of public distributed programs.

The questions for public discussion by interested organizations are intended to represent on special web site.

The profile of electronic library as a total combination of standards on program interfaces and protocols may be constructed with the conceptual model of EL.

This model is offered as a expansion of basic open systems environment / reference model (OSE/RM) for the area of applications to describe the Application Program Interface (APIs).

Opinions of the interested organizations, which will be collected on a site by way of discussion of the questions put above, it is supposed to base on formation of a national information infrastructure of Russia for science, culture and education.

1. Введение

Настоящая работа выполнена при поддержке Российского Фонда фундаментальных исследований (гранты 01-01-00947 и 07-100 07-900).

В докладе авторов на Третьей Всероссийской конференции по электронным библиотекам, г. Петрозаводск, 11-13 сентября 2001 г. [1] были предложены объекты стандартизации в профилях электронных библиотек России.

Функции электронных библиотек (ЭБ) реализуются с помощью информационных систем (ИС), которые поддерживают электронные представления информационных ресурсов. К информационным ресурсам ЭБ относятся:

1. каталоги и метакаталоги;
2. библиографические описания и тематические подборки;
3. полнотекстовые документы (книги, периодические издания и отдельные статьи из них, диссертации, рефераты, депонированные отчеты и рукописи, географические карты, патенты и стандарты, ноты, музыкальные записи и т.д.);
4. свободно распространяемое программное обеспечение.

В числе объектов стандартизации в предыдущем докладе были названы:

1. форматы представления электронных каталогов;
2. услуги поиска в онлайн-овых публичных каталогах и библиографических базах данных;
3. услуги межбиблиотечного обмена полными текстами документов;
4. форматы представления и хранения экземпляров электронных коллекций.

Кроме того, построение ИС, реализующих функции ЭБ, должно опираться на стандартные интерфейсы прикладного программирования, которые определяют взаимодействие между прикладными программами ЭБ и программно-аппаратной средой, где они выполняются.

Рекомендации по стандартам для российских ЭБ следует ориентировать на международные стандарты, принятые мировым библиотечным сообществом. А совокупность рекомендуемых стандартов целесообразно представлять в виде профилей ЭБ согласно методологии функциональной стандартизации в области информационных технологий [2, 3, 4].

Для конкретизации этой методологии применительно к области ЭБ в настоящем докладе в порядке обсуждения предлагается концептуальная модель, представляющая собой расширение эталонной модели среды открытых систем OSE/RM по ГОСТ ИСО/МЭК ТО 10000-3-99 на функциональную область приложений ЭБ, функционирующих поверх системного ПО и ПО промежуточного слоя.

Вопросы выбора стандартов для названных выше объектов стандартизации российских ЭБ, а также методологические вопросы построения профилей ЭБ нуждаются во всестороннем обсуждении с участием заинтересованных организаций. Такое обсуждение предполагается организовать на специальном сайте в сети Интернет (www.elbib.ru).

В настоящем докладе даются комментарии к вопросам, которые предполагается поставить на обсуждение.

II. Варианты выбора стандартов на форматы представления электронных каталогов и метакаталогов.

Стандартный формат обмена библиографическими данными MARC (Machine Readable Cataloguing) является основой для построения электронных каталогов библиотек и библиографических баз данных. Стандарты MARC поддерживаются техническим комитетом ISO TC 46 / SC 4, международной ассоциацией библиотек IFLA, Библиотекой Конгресса США, Британской библиотекой, Национальной библиотекой Канады. Исторически сложились разные версии формата MARC: американская версия US-MARC, разработанная для Библиотеки Конгресса США, британская версия UKMARC, разработанная для Британской библиотеки. Позже в 1996 г. был разработан международный формат UNIMARC, который поддерживается Постоянным Комитетом UNIMARC Ассоциации IFLA [5].

В 1998 г. Российской национальной библиотекой был разработан российский вариант формата UNIMARC [6].

Формат UNIMARC недавно выбран многими европейскими библиотеками для электронного обмена каталожными данными, а также Консорциумом европейских исследовательских библиотек (CERL) для формирования библиографической базы данных, объединяющей каталожные данные из разных библиотек Европы.

Для построения базы данных онлайн-компьютерного центра OCLC (Online Computer Library Center) применен формат USMARC.

Эта база данных содержит более 30 млн. библиографических записей и используется более чем 20 тыс. библиотеками в 61 стране. Ежедневно в нее добавляются более 34 тысяч записей по монографиям, периодическим изданиям, аудио и видеоматериалам, картам, архивам, рукописям, компьютерным файлам. Эти записи оформляются в соответствии с англо-американскими правилами каталогизации (AACR 2), принятым в 1978 г. и дополненными в 1993 г [7]. База данных OCLC является результатом кооперированной деятельности по каталогизации Библиотеки конгресса США по медицине, Национальной библиотеки Канады и Национальной библиотеки Австралии. Обеспечение доступа российских библиотек к OCLC, как к мировой базе библиографических данных, представля-

ется чрезвычайно важным с точки зрения интеграции с мировым библиотечным сообществом.

В связи с изложенным выше необходимо выдвинуть для обсуждения среди заинтересованных организаций России следующие вопросы:

1. какой выбрать стандартный формат построения электронных каталогов для библиотек России: UNIMARC, USMARC или RUSMARC?

2. является ли формат RUSMARC строгим подмножеством формата UNIMARC, или же при его применении потребуются конвертеры, чтобы обеспечить взаимный обмен каталожными данными с мировым библиотечным сообществом?

3. что необходимо предпринять для обеспечения доступа российских библиотек к базе данных OCLC в технологическом и организационном плане?

III. Варианты выбора стандартов на форматы электронных сообщений, связанных с межбиблиотечным обменом.

Услуги межбиблиотечного обмена и спецификации протоколов обмена определены международными стандартами ISO 10160: 1997 и ISO 10161: 1997, поддерживаемыми Техническим комитетом ISO TC 46 / SC4. Стандарты межбиблиотечного обмена (ILL – InterLibrary Loan) определяют спецификации электронных сообщений и протоколы обмена, касающиеся запросов на требуемые читателями информационные ресурсы библиотек, ответов на эти запросы, сведений об отсылке затребованных читателями изданий и т.д. Для использования двух указанных международных стандартов существуют международные профили, основанные на семиуровневой эталонной модели ВОО (OSI-ISO). Часть 4 профиля ALD 22 специфицирует применение электронной почты по протоколу X.400 для обмена сообщениями. Однако указанные стандарты и профили не получили широкого применения в библиотечной практике, поскольку гораздо практичнее оказалось использовать для межбиблиотечного обмена услуги сети Интернет с ее стеком протоколов TCP/IP, а не протоколы ВОО (OSI), учитывая, что существующие реализации стека протоколов OSI громоздки и не обеспечивают функционирование приложений на персональных компьютерах.

Такие же соображения могут быть приведены и в отношении услуг поиска информации в соответствии со стандартами ISO (см. ниже).

С учетом изложенного для обсуждения с заинтересованными организациями предлагаются следующие вопросы:

1. необходимость вообще стандартизовать услуги межбиблиотечного обмена для библиотек России и соответственно необходимые форматы электронных сообщений и протоколы;

2. если стандартизовать услуги межбиблиотечного обмена, то какие протоколы электронной почты, стандартно доступные в Интернет, следует выбрать: MIME, S/MIME?

3. следует ли при этом учитывать проекты в области электронного межбиблиотечного обмена, поддерживаемые в настоящее время Европейской комиссией (DALI, UNIVERSE, ION OSI, AIDA)?

IV. Варианты выбора стандартов на услуги поиска информации в онлайн-овых публичных каталогах и библиографических базах данных.

Услуги, команды и протоколы для поиска информации в библиографических базах данных определены международными стандартами ISO 8777: 1993, ISO 10162: 1993, ISO 10163: 1993. Они базируются на спецификациях протоколов прикладного (седьмого) уровня эталонной модели ВОС (OSI). Международные профили OSI-ISP были стандартизованы в ИСО. Однако широкого применения в библиотечном деле эти стандарты и профили не нашли.

Более перспективным представляется использование для услуг поиска информации протокола Z 39.50, стандартизованного в ИСО (стандарт ISO 23950). В США протокол Z 39.50 составляет основу для услуг Government Information Locator Service (GILS). В Европе на протоколе Z 39.50 базируется проект онлайн-овых публичных каталогов OPAC Network in Europe (ONE). В рамках Европейской библиотечной программы протокол Z 39.50 реализуется в проектах SOCKER и ARCA (Access to Remote Catalogues).

В связи с тем, что в большинстве случаев телекоммуникационная среда опирается на стек протоколов TCP/IP сети Интернет, в Европе разрабатывается проект Paragon, финансируемый Европейским сообществом.

Одновременно Европейское сообщество финансирует проект Europe Gate, в котором обеспечивается шлюз между протоколом Z 39.50 и Web с услугами электронной почты, доступными в Интернет, и услугами протокола HTTP.

Реализации протокола Z 39.50 разработаны и в России, например в ОИГГ и СО РАН, г. Новосибирске [8, 9], в частности, информационная система, состоящая из сервера Z 39.50 ZooPARC и сервера баз данных.

Поскольку услуга поиска информации ISO-SR поверх стека протоколов OSI в соответствии со стандартами ISO не получила широкого распространения, и в качестве ее следующей версии рассматривается также протокол Z 39.50, предлагается принять для электронных библиотек России стандартный протокол Z 39.50.

Для обсуждения с заинтересованными организациями предлагаются следующие вопросы:

1. достаточно ли указания в профиле ЭБ для российских библиотек только стандарта Z 39.50?

2. потребуются ли для российских библиотек шлюзы (мосты) между реализациями протокола Z 39.50 поверх стека протоколов TCP/IP и реализациями услуги поиска ISO-SR поверх стека протоколов BOC, подобно тому, как это предусмотрено в Европейском проекте Europe Gate?

3. какие реализации протокола Z 39.50 предпочтительно выбрать из числа уже разработанных?

V. Варианты выбора стандартов представления информационных ресурсов электронных библиотек и форматов электронного обмена полными текстами документов.

Форматы и протоколы электронного обмена полными текстами документов определены соглашениями GEDI (Group on Electronic Data Interchange). Третья версия этих соглашений, принятая в 1997 г. определяет в качестве рекомендуемых форматов: TIFF, PDF, JPEG, PostScript с механизмом обмена по протоколу электронной почты Интернет MIME.

Соглашения GEDI широко используются в Европе, США и Австралии для поддержки национальных и международных служб доставки документов в библиотечном мире. В Европе соглашения GEDI с учетом многонационального аспекта европейских стран реализуются в проекте EDIL, финансируемом Европейским сообществом.

Информационные ресурсы электронных библиотек могут быть представлены в виде:

1. полных текстов документов, возможно, содержащих иллюстрации;
2. векторных графических представлений;
3. растровых изображений;
4. аудиоматериалов;
5. видеоматериалов;
6. текстов свободно распространяемых программ.

Для каждого из этих представлений в профиле ЭБ требуется указать выбранные стандарты услуг из числа вариантов рассмотренных выше.

Сводное предложение в виде таблицы для рассмотрения заинтересованными организациями будет опубликовано на сайте www.elbib.ru.

VI. Концептуальная модель для построения профиля электронной библиотеки.

Для формирования профилей электронных библиотек авторами предлагается концептуальная модель, представляющая собой расширение эталонной модели среды открытых систем OSE/RM по ГОСТ Р ИСО/МЭК

ТО 10000-3-99. В предлагаемой концептуальной модели услуги, реализуемые прикладным программным обеспечением (приложениями) ИС ЭБ, представлены по тем же четырем функциональным группам, что и услуги среды. Между слоем приложений и слоем услуг среды, включающим системное ПО и ПО промежуточного слоя, должны быть указаны стандартные интерфейсы прикладного программирования.

VII. Заключение

Мнения заинтересованных организаций, которые будут собраны на сайте в порядке обсуждения поставленных выше вопросов, предполагается положить в основу формирования национальной информационной инфраструктуры России для области науки, культуры и образования.

Литература

1. Е.Н.Филинов, А.В.Бойченко. Объекты стандартизации в профилях электронных библиотек России // Сборник трудов Третьей Всероссийской конференции по электронным библиотекам. Петрозаводск, 11-13 сентября 2001 г.
2. ГОСТ Р ИСО/МЭК ТО 10000-1-99. Информационная технология. Основы и таксономия функциональных стандартов. Часть 1. Основные положения и основы документирования.
3. ГОСТ Р ИСО/МЭК ТО 10000-2-99. Часть 2. Принципы и таксономия профилей ВОС.
4. ГОСТ Р ИСО/МЭК ТО 10000-3-99. Часть 3. Принципы и таксономия профилей среды открытых систем.
5. UNIMARC. www.ifla.org/VI/3/p1996-1/unimarc.htm
6. Российский коммуникативный формат библиографических записей в машиночитаемой форме (Российский вариант UNIMARC). Изд. РНБ. Санкт-Петербург. 1998 г.
<http://www.rba.ru:8101/rusmarc/rusmarc/format1.htm>
7. The Anglo-American Cataloging Rules (1994).
8. Ottawa: Canadian Library Association
9. London: Library Association Publishing;
10. Chicago: American Library Association.
11. О.Л.Жижимов, С.В.Скибин. О доступе к реляционным СУБД по протоколу Z.39.50 // Сборник трудов третьей Всероссийской конференции по электронным библиотекам. Петрозаводск, 11-13 сентября 2001 г.
12. О.Л.Жижимов. Введение в Z.39.50.
<http://geolibr.uiggm.nsc.ru/does/z39.50/z-intro>