

# Электронная библиотека по научному наследию как фактографическая система

© В.Б.Барахнин    © А.М.Федотов    © О.А.Федотова  
Институт вычислительных технологий СО РАН,  
Государственная публичная научно-техническая библиотека СО РАН,  
Новосибирский государственный университет  
[bar@ict.nsc.ru](mailto:bar@ict.nsc.ru)    [fedotov@sbras.ru](mailto:fedotov@sbras.ru)    [o4f8@mail.ru](mailto:o4f8@mail.ru)

## Аннотация

Работа посвящена обсуждению модели электронной библиотеки (ЭБ) по научному наследию и ее архитектуре. Особое внимание уделено функциональным требованиям к ЭБ такого типа, которые определяются, во-первых, информационными потребностями исследователей в фактографической информации, а во-вторых, обеспечением надежного и долговременного хранения информации.

Работа выполнена при частичной поддержке РФФИ (проекты 11-07-00561, 12-07-00472, 13-07-00258), президентской программы «Ведущие научные школы РФ» (грант НШ 6293.2012.9) и интеграционных проектов СО РАН.

## 1 Введение

Научное наследие – это опубликованные результаты научных исследований и экспериментов, библиографические и фактографические базы данных, сведения об ученых, их научной деятельности, публикациях, проектах и т.п., а также большое количество неопубликованных документов, таких как отчеты, письма, воспоминания, записки, фотоматериалы и т.п. Эти ресурсы представляют большой интерес для научного сообщества и представителей общественности.

В классической монографии [1], изданной ВИНТИ еще в 1976 году, выделено два типа информационных потребностей исследователя: потребности в сведениях об источниках необходимой научной информации и потребности в самой необходимой научной информации (фактах). Для удовлетворения информационных потребностей первого типа предназначены информационные

системы (ИС), получившие название «документальных» (библиографических), второго типа – «фактографических». В настоящее время наиболее востребованным средством информационного обеспечения научной деятельности становятся интеллектуальные системы, сочетающие возможности информационных систем обоих названных типов и позволяющие удовлетворять информационные потребности квалифицированного пользователя в соответствии со схемой «документ – факт – рассуждение» [2].

В интеллектуальных информационных системах в качестве составного компонента выступают рассуждающая система, формализующая правила логического вывода, и интеллектуальный интерфейс.

Для уточнения смысла, вкладываемого в термин «факт» применительно к информационным системам по научному наследию, представляется целесообразным использование семиотического подхода, подобно тому, как это было сделано в работе [3] для терминов «информация», «знание», «тезаурус», «онтология». В этой работе, в частности, показано, что данные соответствуют синтаксическому уровню сообщения (в том числе документа), информация (в узком смысле!) – семантическому, а знания – прагматическому. Отсюда вытекает, что функционирование интеллектуальной информационной системы основано на двух противоположных процессах: при пополнении ИС новыми сведениями происходит преобразование семантической информации в данные, однако непосредственно потребности пользователя удовлетворяет обратный процесс – извлечение из данных нужной пользователю информации и знаний.

Ввиду того, что информация в ИС отображает некоторые сущности (предметы, процессы, явления, персоны, публикации, факты, ключевые термины и т.п.), следует рассматривать информационную систему как множество информационных объектов – наборов данных, представляющих (описывающих) эти сущности в ИС.

Эффективным средством описания информационных объектов в ИС являются метаданные – данные, являющиеся неотъемлемой частью информационного объекта и описывающие реальный объект или группу объектов (рис.1).

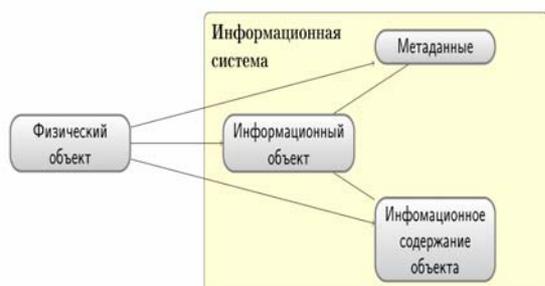


Рис.1. Структура информационной системы.

Для информационных систем по научному наследию важной проблемой является идентификация ресурсов, определяющая конкретно для каждого факта, кто имеет к нему отношение (например, является его автором), где и когда он имел место, с какими другими фактами и объектами он связан. Для этого необходима поддержка различных уровней абстракции при создании метаописаний (метаданных) информационных объектов: от кратких описаний до очень подробных. Далее будем использовать следующее понимание факта: «входящая в текст документа характеристика сущности, описываемой в онтологии информационной системы, представляемая как единичное значение данных». Факт может быть извлечен из информационного содержания объекта, либо определен экспертом. Факт может определять как свойства (атрибуты) объекта, так и его связь с другими объектами.

Для поддержки сложных функций поиска и классификации информации недостаточно хранить только полнотекстовые описания. Необходимы поддержка поиска по атрибутам, полнотекстового поиска, а также просмотр ресурсов по категориям и словарям-классификаторам. При этом выбор классификаторов определяется степенью специализации системы. Так, в электронной библиотеке «Научное наследие России» [4], создаваемой ведущими библиотеками и другими организациями РАН (БАН, БЕН, ИНИОН) в рамках технологий Единого научного информационного пространства (ЕНИП) РАН, разрабатываемых ВЦ РАН, используется классификатор ГРНТИ. Этот классификатор имеет весьма общий характер, но он вполне отвечает потребностям пользователей рассматриваемой ЭБ (по крайней мере, на данном этапе), поскольку в настоящее время (август 2013 г.) к почти каждому разделу верхнего уровня ГРНТИ отнесено менее 1 тысячи ресурсов (исключение составляют разделы «Философия» и «Религия. Атеизм», но и в них количество ресурсов пока не превосходит 2 тысяч). Если же ЭБ содержит большое количество ресурсов, относящихся к той

или иной конкретной области науки, то требуется использование соответствующего специализированного классификатора. Например, в математических ЭБ обычно используется Классификатор математических сущностей MSC2010 [5] (или его более ранняя версия MSC2000).

В большинстве существующих ИС документы являются слабо структурированными: хотя и снабженными метаданными, но содержащими неструктурированные элементы. Поэтому актуальной задачей является разработка теоретических основ и моделей создания ИС, способных в автоматизированном режиме извлекать метаданные и факты из электронных документов достаточно произвольной структуры. Ее решение позволит получать новую информацию и знания [2].

В настоящий момент значительная часть информационных ресурсов по научному наследию хотя и переведена в цифровую форму, но недоступна широкому кругу научной общественности, а ресурсы, представленные в Интернет, разрознены, недостаточно систематизированы и структурированы. При создании их описаний недостаточное внимание уделяется вопросам интероперабельности: слабо применяются соглашения и рекомендации по стандартизации представления документов и средства интеграции разнородных информационных ресурсов. Под интероперабельностью ИС понимается степень ее способности взаимодействовать с другими ИС, в том числе и с человеком. Но если при взаимодействии с человеком (как с информационной системой) основная нагрузка на обеспечение взаимопонимания ложится на человека, который в состоянии обработать даже плохо организованную информацию, то для обеспечения эффективного взаимодействия между собственно информационными системами требуются специальные технологические методы и общие соглашения. Это приводит к требованию соответствия всех схем данных, интерфейсов и протоколов соответствующим международным стандартам и рекомендациям [6, 7].

## 2 Определение электронной библиотеки

Проблема поиска информации – одна из вечных проблем человечества. Чтобы решить проблему доступа к информации, человечество создало библиотеки – универсальную систему хранения, систематизации и каталогизации информации и знаний.

Электронная библиотека (ЭБ) – это структурированная каталогизированная коллекция разнородных электронных документов, снабженная средствами навигации и поиска (в отличие от печатных изданий, микрофильмов и других носителей). ЭБ способна не только обеспечить многосторонний поиск в каталоге, но и

предоставить пользователю непосредственно найденный ресурс (публикацию, документ, фотографию, описание факта и др.), а также дополнительные сведения о нем, например, информацию об авторах, библиографию, организации и т. п.

За высокой популярностью слов «электронная библиотека» стоит не только и не столько дань моде, сколько попытка охарактеризовать новый феномен – возникновение принципиально нового класса систем, призванных аккумулировать и распространять информацию в электронной форме. А большой интерес к самим системам данного класса объясняется потребностями общества и наличием развивающихся возможностей по их удовлетворению. В связи с этим можно сформулировать основные цели, стоящие перед ЭБ:

- обеспечение доступа к информации;
- сохранение научного и культурного наследия;
- повышение эффективности научных исследований и обучения.

В существующих разработках ЭБ, как правило, поиск и доступ к информации обеспечиваются только посредством визуальных графических интерфейсов. Это хорошо для пользователя-человека, но не годится для пользователя-системы. Для обеспечения функций поиска вне графических интерфейсов требуется поддержка специальных сетевых сервисов и языков запросов. В идеальном случае все ИС должны поддерживать единый поисковый профиль и единый язык запросов.

Однако в общем случае под словосочетанием «электронная библиотека» могут фигурировать совершенно различные объекты, такие как архивы цифрового контента и наборы программного обеспечения для управления этим контентом. Электронной библиотекой может называться система сетевых сервисов, предоставляющих доступ к цифровому контенту, объединенных единой системой управления этим доступом [8]. Такое определение ЭБ полностью соответствует определению традиционной библиотеки как организации в системе, например, Министерства культуры [6]. Разумеется, мы не претендуем на то, чтобы считать это определение «наиболее правильным», однако именно оно оказывается наиболее функциональным при работах по сохранению научного наследия в виде электронных документов.

В настоящее время нет какой-либо универсальной системы поддержки ЭБ, которая отвечала бы всем требованиям и ожиданиям пользователей. Анализ существующих систем ЭБ проведенный, например, в [9], показывает их разнородность на нескольких уровнях:

- на уровне информационной модели, которую они обеспечивают;
- на уровне поддержки пользователей и групп пользователей;

- на уровне функциональных возможностей.

Из-за этой разнородности и игнорирования нужд пользователей возникает ряд проблем:

- интеграция информации из различных ЭБ;
- сравнение ЭБ по предоставляемой функциональности;
- оценка и сравнение производительности различных систем ЭБ;
- добавление новых типов хранимых объектов;
- добавление новых функциональных возможностей;
- резервное копирование.

В настоящее время существуют достаточно мощные ИС, удовлетворяющие в той или иной степени потребности научных работников в информации, однако основной недостаток большинства систем – ограниченность возможностей обеспечения интеграции ресурсов как внутри каждой из систем, так и с внешними системами. Основу разработки ЭБ составляют стандарты и международные рекомендации, формирующие профиль ЭБ, под которым понимается набор из одного или нескольких базовых нормативно-технических документов (стандартов и спецификаций), ориентированных на решение определенной задачи (реализацию заданной функции либо группы функций приложения или среды) с указанием, при необходимости, выбранных классов, подмножеств, опций базовых стандартов, которые являются необходимыми для выполнения конкретной функции [10]. Наиболее важным является профиль метаданных информации, циркулирующей в системе. Выбор профиля должен основываться на выполнении следующих требований:

- включать в себя основные типы информации, требующейся для поддержки научной работы;
- быть открытым, т. е. обеспечивать доступ к соответствующей информации по этим описаниям;
- быть расширяемым, т. е. обеспечивать возможность детализации описаний;
- обеспечивать возможности интеграции информации;
- обеспечивать возможности уникальной идентификации информации;
- обеспечивать возможности размещения и поиска информации в распределенной среде;
- быть ориентированным на современные и перспективные технологии описания и использования информации;
- обеспечивать возможность интероперабельности с внешней средой.

При работе с цифровыми объектами человечество уже выработало определенный набор стереотипов, отсутствие которых вызывает дискомфорт [6]. Одним из элементов этого набора является требование наличия взаимных ссылок между цифровыми объектами, проявляющихся,

например, в виде гиперсвязей в пользовательских графических интерфейсах просмотра информации. Реализация взаимных ссылок в цифровых документах не представляет большой сложности, однако при этом проявляются специфические моменты. Во-первых, электронный объект с реализованными связями уже не совсем соответствует своему печатному оригиналу. Во-вторых, внедренные в объект связи должны быть гарантировано актуальными. Так появляется требование обеспечения ссылочной целостности данных. Это очень жесткое требование, которое трудно обеспечить даже в хорошо формализованных системах управления базами данных. Результат □ новый цифровой объект как самосогласованное хранилище цифрового контента, или база данных цифровых объектов.

С другой стороны, в ЭБ объекты хранения могут содержать информацию, которая не имеет к объектам хранения традиционных библиотек вообще никакого отношения. Речь идет об электронных копиях элементов хранения традиционных архивов, о видео- или аудиоинформации, полученной разными способами, о научных или других фактах и т. п.

Существует достаточно много технологических разработок информационных систем для электронных библиотек, так или иначе ориентированных на поддержку научных исследований. Среди них следует отметить информационные системы, близкие к фактографическим, например ИСИР (ЕНИП) РАН [11], ИРИС СО РАН [8], euroCRIS [12], и документальные, например eLibrary [13], Информика [14], MathNET [15]. Названные системы в той или иной степени удовлетворяют потребности исследователей в информации, однако каждая из них имеет определенные недостатки (см., например, [2]).

## **2 Функциональные требования к модели электронной библиотеки по научному наследию**

Как уже отмечалось выше, основными целями создания ЭБ по научному наследию являются:

- предоставление научным работникам быстрого доступа к информационным ресурсам по научному наследию;
- предоставление результатов фундаментальных научных исследований мировому сообществу;
- предотвращение утраты ценных научных коллекций для будущих поколений ученых;
- создание новых технологий научных исследований, эффективного инструментария для их проведения.

Как известно большая часть научной информации быстро устаревает. Но это не относится к материалам по научному наследию. Для этого

типа информационных ресурсов важно хранить описание жизненного цикла этих ресурсов и иметь возможность восстановить состояние ресурса на любой момент времени. Кроме того, существуют информационные ресурсы, которые могут быть доступны длительное время. К таковым, например, относятся документы, имеющие длительную юридическую силу, патенты, мультимедийная информация об исторических событиях, которая может быть востребована через любой период времени. Кроме того, научные отчеты институтов, речи ученых, письма и служебные записки могут также иметь огромную историческую значимость, становясь более ценной со временем. Поэтому ЭБ должна поддерживать возможность длительного хранения информационных ресурсов с возможностью восстановления их.

Для документов по научному наследию важной проблемой является идентификация информационных ресурсов [16], определяющая конкретно для каждого факта, кто является его автором, где и когда он получен, с какими другими фактами он связан. Для этого необходима поддержка различных уровней абстракции при описании информации от кратких описаний, до очень подробных описаний информационных объектов.

Исходя из целей ЭБ по научному наследию и анализа существующих систем, направленных на поддержку научных исследований, можно сформулировать следующие функциональные требования к модели ЭБ по научному наследию:

- надежное долговременное и защищенное от исчезновения хранение информации;
- актуальность, полнота, достоверность происхождения документов;
- историчность информации;
- географическая привязка информации;
- наличие большого числа словарей-классификаторов (справочников), для обеспечения идентификации и классификации ресурсов;
- поддержка неоднородных и слабо структурированных информационных ресурсов;
- поддержка взаимосвязей информационных ресурсов;
- предоставление информации пользователю в виде, выбранном пользователем;
- наличие интеллектуальных служб обслуживания запросов пользователя;
- наличие программных интерфейсов для поддержки аналитической работы пользователя с помощью программных приложений;
- поддержка требований интероперабельности как на программном, так и на семантическом уровне;
- поддержка работы с внешними источниками.

Наиболее важным выводом из вышесказанного является то, что информационная модель ЭБ должна быть многоуровневой и состоять как минимум из

следующих компонент [17]: хранилище данных – репозиторий, сервер метаданных, сервер приложений, словари-справочники (см. рис.2).



Рис. 2 Архитектура электронной библиотеки.

### 3 Выбор метаданных для ЭБ по научному наследию

В работах [7, 16] был определен профиль ЭБ как необходимый набор стандартов и компонентов информационной системы, ориентированной на научные исследования.

В настоящее время существует большое количество систем метаданных, предназначенных для описания различных классов информационных объектов. Использование систем метаданных (схем данных) пока еще недостаточно формализовано, что убедительно показано в обзоре [18]. Информационные системы, ориентированные на одинаковые классы информационных объектов, используют различные, часто оригинальные схемы и форматы метаданных, а также разные подходы к решению прикладных задач. Решением этих проблем занимаются многие организации во всем мире, например W3C, DCMI, OCLC, IFLA, IETF, ISO.

Метаданные необходимы для решения следующих задач:

- предоставление сведений об объекте для получения представления о его содержании, структуре, способах использования и т. д.;
- сбор и систематизация информации об объектах описания;
- выбор из множества объектов определенного подмножества по формальным признакам и сопоставление объектов по формальным признакам;
- внутрисистемные технологические задачи, связанные с обеспечением подготовки объектов, размещением объектов в информационном фонде и т. п.;
- внешние технологические задачи, связанные, прежде всего, с обменом данными с внешними информационными системами.

Таким образом, в соответствии с [18], метаданные целесообразно рассматривать как особого рода информационные ресурсы, выполняющие в использующих их ИС весьма разнообразные функции.

Как мы уже отмечали, основу содержания электронной библиотеки по научному наследию составляют информационные объекты, которые представляют следующие основные типы сущностей [7, 16, 17]:

- субъекты: персоны, организации и т. п.;
- объекты – единицы хранения: публикация, документ, факт, научный результат, мероприятие, фотография и др.;
- отношения: понятие, ключевой термин, событие, время, место.

В отличие от общепринятых документных (библиографических) ЭБ указание на субъекты дается ссылкой на экземпляр сущности субъект, что позволяет корректно решать задачу идентификации объектов. Схема отношений в ЭБ по научному наследию является персоноцентричной: все объекты и отношения, понятия, факты, мероприятия, публикации и др. жестко привязываются к персонам.

Используемый профиль определяет список элементов данных (полей), необходимых для создания записи соответствующего типа и раскрывает содержание элементов данных [6, 7]. Для эффективной работы сервера приложений необходимо использовать набор словарей-классификаторов, содержащих как классификационные признаки, так и наборы ключевых терминов (с отношениями порядка), по которым производится систематизация и классификация материала.

Для формирования метаданных применяются несколько стандартов, являющихся расширениями рекомендаций Dublin Core [19] и Qualified Dublin Core (QDC). Для документов нами была расширена стандартная схема метаданных QDC полями, включающими основные требования государственного стандарта МЕКОФ [20].

Словари (ключевые признаки, ключевые термины) – это особый вид метаданных, которые отражают наиболее существенные свойства объекта, имеющие наибольшее значение с точки зрения ИС, и их специфика определяются терминологией конкретной предметной области, которой посвящена ЭБ. Необходимо рассматривать различные типы ключевых терминов, а именно:

- ключевые термины в стандартном понимании;
- ключевые термины, описывающие персону;
- ключевые термины, описывающие организацию;
- ключевые термины, описывающие временные периоды;
- ключевые термины, описывающие географические понятия,

а также тематические словари-классификаторы, тезаурусы, описания предметной области данной

научной школы и классификаторы документов в соответствии с МЕКОФ.

Имеется ряд российских (например УДК, ГРНТИ) и международных (например MSC2010 [5], ORTELIUS [21]) словарей для классификации научных данных. Однако в целом эти словари содержат только общенаучную информацию и не годятся для систематизации материалов по научному наследию конкретной научной школы.

Метаданные существенным образом зависят от природы и структуры объектов реального мира, от способа представления их в виде информационных объектов и от специфики ИС. Учитывая это, необходимо классифицировать описываемые объекты. Законченная совокупность правил, достаточная для формирования метаданных в определенном классе ИС и (или) для решения определенного класса задач над информационными объектами представляет собой систему метаданных.

Функционирование ИС связано с разнообразными процессами по созданию метаданных, их модификации, проверке корректности, предоставлению метаданных конечному пользователю и решению прикладных задач. Все эти процессы являются взаимосвязанными, их выполнение усложняется, как правило, большим количеством объектов, на представление и работу с которыми направлена ИС. Реализация этих процессов и управление ими требуют специальных средств и методов, которые в совокупности с метаданными можно рассматривать как отдельную подсистему – систему метаинформационного сопровождения.

#### 4 Практическая реализация

Рассмотренная модель информационной системы, работающей с материалами научного наследия, реализуется на примере научной школы Алексея Андреевича Ляпунова – основателя теоретического программирования и отечественной кибернетики.

Основной каталог информационных ресурсов сервера метаданных информационной системы строится в соответствии со схемой метаданных МЕКОФ. Для долговременного хранения документов использовался репозиторий DSpace. Стандартная схема метаданных DSpace была расширена полями, отвечающими основным требованиям МЕКОФ. Для поддержки процесса наполнения полнотекстовых баз созданные профили метаданных были зарегистрированы в системе DSpace и в соответствии с ними были настроены рабочие процессы, а также пользовательский интерфейс системы. Для того чтобы выполнять обмен метаданными между DSpace в соответствии с расширенным профилем, был создан сервис, выполняющий преобразование схем метаданных из внутренней схемы DSpace в схему сервера метаданных и в схему Dublin Core с использованием квалификаторов. Реализован также OAI-сервис [22],

который в пакетном режиме периодически, в соответствии с расписанием, проводит синхронизацию метаданных репозитория и сервера метаданных. Для заполнения основного каталога метаданных в соответствии с основными схемами метаданных используется контролируемые словари из справочного блока сопровождения.

#### 5 Заключение

Разработанная модель информационной системы может быть использована как типовая модель системы для работы с документами, связанных с научным наследием, поскольку решает основные задачи, предъявляемые к этим системам: обеспечение системы надежного долговременного хранения цифровых (электронных) документов с сохранением всех смысловых и функциональных характеристик исходных документов; обеспечение «прозрачного» поиска и доступа пользователей к документам, как для ознакомления, так и для анализа содержащихся в них фактов; организация сбора информации по удаленным ЦД, поддерживающих протокол OAI.

#### Литература

- [1] А.И.Михайлов, А.И.Черный, Р.С.Гиляревский. Научные коммуникации и информатика. М: Наука, 1976.
- [2] Ю.И.Шокин, А.М.Федотов, В.Б.Баракнин. Проблемы поиска информации. Новосибирск: Наука, 2010.
- [3] В.Б. Баракнин, А.М.Федотов. Уточнение терминологии, используемой при описании интеллектуальных информационных систем, на основе семиотического подхода. Известия вузов. Проблемы полиграфии и издательского дела. 2008. № 6. С.73-81.
- [4] Электронная библиотека «Научное наследие России». <http://nasledie.enip.ras.ru/>
- [5] Классификатор математических сущностей MSC2010. <http://www.ams.org/mathscinet/msc/msc2010.html>
- [6] О.Л.Жижимов, Н.А.Мазов, А.М.Федотов. Некоторые заметки об эволюции цифровых репозитариев традиционных библиотек к полнофункциональным электронным библиотекам. Вестник Владивостокского государственного университета экономики и сервиса. Территория новых возможностей. 2010. Т 7. № 3. С. 55-63.
- [7] А.М.Федотов, В.Б.Баракнин, О.Л.Жижимов, О.А.Федотова. Технология создания корпоративных информационных систем учета трудов научных работников. Вестник Новосибирского государственного университета. Серия: Информационные технологии. 2011. Т. 9. Вып. 2. С. 31-41.
- [8] Ю.И.Шокин, А.М.Федотов, О.Л.Жижимов, А.Е.Гуськов, С.В.Столяров. Электронные

- библиотеки – путь интеграции информационных ресурсов Сибирского отделения РАН. Вестник КазНУ, специальный выпуск. Алма-Ата, Казахстан, Казахский национальный университет им. аль-Фараби. 2005. № 2. С. 115-127.
- [9] L.Candela, D.Castelli, N.Fuhr, Y.Ioannidis, С.-P.Klas, P.Pagano, S.Ross, C.Saidis, H.-J.Schek, H.Schuldt, M.Springmann. Current Digital Library Systems: User Requirements vs Provided Functionality. IST-2002-2.3.1.12. Technology-enhanced Learning and Access to Cultural Heritage. March 2006.
- [10] ГОСТ Р ИСО / МЭК ТО 10000-2-99. Информационная технология. Основы и таксономия функциональных стандартов. Часть 2. Принципы и таксономия профилей ВОС.
- [11] А.Н.Бездушный, А.А.Бездушный, В.А.Серебряков, В.И.Филиппов. Интеграция метаданных Единого Научного Информационного Пространства РАН. М.: Вычислительный Центр им. А.А.Дородницына РАН, 2006.
- [12] Информационная система EuroCRIS. <http://www.eurocris.org/>
- [13] Научная электронная библиотека eLIBRARY.RU. <http://elibrary.ru/>
- [14] Государственный НИИ информационных технологий и телекоммуникаций «Информика». <http://www.informika.ru/>
- [15] Общероссийский математический портал Math-Net.Ru. <http://www.mathnet.ru/>
- [16] А.М.Федотов, В.Б.Барахнин, О.Л.Жижимов, О.А.Федотова. Проблемы создания информационных систем учета трудов научных сотрудников СО РАН. Труды IV Международной конференции «Системный анализ и информационные технологии» (САИТ-2011) (Абзаково, Россия, 17-23 августа 2011). Челябинск: ЧелГУ, 2011. С. 85 -91.
- [17] О.Л.Жижимов, А.М.Федотов, О.А.Федотова. Построение типовой модели информационной системы для работы с документами по научному наследию. Вестник Новосибирского государственного университета. Серия: Информационные технологии. 2012. Т. 10. Вып. 2. С. 5-14.
- [18] М.Р.Коголовский. Метаданные, их свойства, функции и классификация. Труды XIV Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» (RCDL'2012). Переславль-Залесский, 15-18 октября 2012 г. С.25-36.
- [19] Core Metadata Initiative. <http://dublincore.org/>
- [20] ГОСТ 7.19-2001 Система стандартов по информации, библиотечному и издательскому делу. Формат для обмена данными. Содержание записи.
- [21] The «Ortelius Thesaurus on Higher Education». <ftp://ftp.cordis.lu/pub/cerif/docs/ortelius.doc>
- [22] The Open Archives Initiative Protocol for Metadata Harvesting. Protocol Version 2.0 of 2002-06-14. Open Archives Initiative: The OAI Executive; OAI Technical Committee. 2004. <http://www.openarchives.org/>

### **A Scientific Heritage Digital Library as a Factographic System**

V.B.Barakhnin, A.M.Fedotov, O.A.Fedotova

The article discusses the model of scientific heritage digital library and architecture of its organization. Particular emphasis is put on functional requirements for such libraries, which are defined, firstly, by the information needs of researchers in factographic information, and secondly, by maintaining long-term reliability of data storage.