

# Структура и реализация электронной библиотеки диссертаций в РГБ \*

Лаврёнова О.А

РГБ  
lavr@rsl.ru

Вежневцев А.А

ВЦ РАН  
filipova@ccas.ru

## Аннотация

Статья посвящена описанию проекта электронной библиотеки диссертаций (ЭБД), который ведет Российская государственная библиотека (РГБ) при поддержке РФФИ. Функции ЭБД поддерживаются следующими компонентами электронного каталога: библиографической базой данных с указанием электронных адресов полных текстов, нормативными файлами для имен лиц и наименований организаций, машиночитаемыми таблицами национальной классификации ББК. В статье рассматриваются следующие аспекты проекта: структура ЭБД; проблемы формирования базы данных и авторское право; технология создания базы данных; поиск в ЭБД, в частности – по темам; краткое сравнение с аналогичными зарубежными проектами; развитие ЭБД на основе программного обеспечения ИСИР РАН, обоснование выбора ИСИР; рассматриваются направления развития системы для ЭБД, приводятся основные результаты.

## 1 Введение

Над проектом электронной библиотеки диссертаций и авторефератов (ЭБД) Российская государственная библиотека (РГБ) работает с 2001г. Цель проекта – формирование полнотекстовой коллекции диссертаций и авторефератов, доступной на Web-сервере РГБ [1] и, частично, в интранет Библиотеки. Изначально предусматривалась дальнейшая возможность создания на базе данной коллекции национального репозитория диссертаций, защищенных в нашей стране.

ЭБД является универсальной как по содержанию, т. е. включает в себя все научные направления, кроме медицины и фармации (до соответствующего соглашения с ЦНМБ), так и по временному показателю - диссертации, включаемых в ЭБ, не ограничиваются хронологическими рамками.

## 2 Текущее состояние проекта.

К настоящему времени в рамках проекта получены следующие результаты:

- Разработаны **главные составляющие проекта**: глоссарий; отчет по предпроектному обследованию объекта автоматизации; функциональные требования к ЭБД; состав ЭБД; правовые основы ЭБД, включая типовый договор с автором; информационное обеспечение системы (порядок комплектования фонда ЭБД); лингвистическое обеспечение (метаданные); требования к программному обеспечению; технология обработки текстов диссертаций и авторефератов для включения в ЭБД и каталогизации электронных документов; набор определителей содержания MARC21 для библиографической записи на диссертацию или автореферат в электронном виде; DTD для XML-разметки текстов диссертаций и авторефератов [2].
- **Организованы процессы** приема диссертаций и авторефератов в электронной форме от авторов, заключения договоров, учета и преобразования файлов для ЭБД (ретроспективная оцифровка фонда диссертаций в рамках данного проекта не проводится).
- **Электронная библиотека диссертаций открыта на Web-сайте РГБ** [1], а также в интранет Библиотеки.
- В **“Бюллетене ВАК”** (2002, №6) опубликовано объявление о приеме диссертаций и авторефератов в ЭБД РГБ.
- **Наш проект РФФИ зарегистрирован на сайте ЮНЕСКО [3]**, который является источником сведений об аналогичных проектах и экспортах.

**Таким образом, с июня 2002 г. электронная библиотека диссертаций (ЭБД) в РГБ работает и имеет официальный статус.**

На 1 июля 2003 г. в ЭБД содержалось 153 полных текста диссертаций и авторефератов, накопленных за год работы, причем с авторами заключено более 250 договоров на передачу файлов в электронную библиотеку. Эти объемы вполне сопоставимы с первыми результатами, полученными в зарубежных проектах.

---

*Труды 5<sup>ой</sup> Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» - RCDL2003, Санкт-Петербург, Россия, 2003.*

Будущая популярность ЭБД среди пользователей была изначально очевидной, в частности, на основе данных о высокой востребованности традиционного фонда диссертаций в РГБ как депозитарного хранилища этого вида документов. Однако цифры, демонстрирующие количество “открытий” полных текстов диссертационных работ пользователями, приведенные на первой странице сайта Открытой русской электронной библиотеки (Open Russian Electronic Library, OREL) РГБ [4], говорят сами за себя: лидер - диссертация по экономике – 1232 открытия, затем идут цифры: 634, 540, 509, 506, 488, 478 и т.д.

### 3 Структура ЭБД

ЭБД является одной из электронных коллекций в рамках электронной библиотеки (ЭБ) **OREL** [4, 5]. Полные тексты диссертаций и авторефератов представлены и выводятся на экран в формате PDF; каждый электронный документ имеет свой электронный титульный лист. Электронные диссертации и авторефераты связываются с метаинформацией, содержащейся в библиографических записях (БЗ) электронного каталога (ЭК) ЭБ РГБ: в поле 856 (формат MARC21) БЗ указываются электронные адреса (URL) PDF-файлов.

Метаданные, описывающие отдельные поля БЗ, содержащие имена лиц, наименования коллективов, индексы и наименования делений российской национальной классификации ББК, коды и наименования специальностей, формируются в виде нормативных и классификационных записей, принятых в библиотечной практике [7, 8].

Можно считать, что в таком виде ЭБД уже эффективно выполняет свои функции. Во всяком случае, **для всех авторов, имеющих электронную версию своей диссертации и/или автореферата и желание поместить их в электронной библиотеке РГБ, обеспечена такая возможность.**

### 4 Проблемы формирования базы данных и авторское право

При всей заинтересованности пользователей в создании общедоступной ЭБД нельзя таким же образом охарактеризовать отношение к ней самих авторов диссертаций. Судя по предпроектному анкетированию [5, 6], авторы, имеющие бесплатную возможность и очень часто желание опубликовать свои работы на Веб-сайте РГБ, по различным причинам, не спешат осуществить этот процесс. Даже заключив договор с РГБ, автор может не присылать файл в течение нескольких месяцев. Для обеспечения запросов своих пользователей Библиотека заинтересована в сканировании и размещении на сайте диссертаций из своего фонда по наиболее спрашиваемым специальностям за счет собственных средств. Однако это стремление наиболее качественно выполнять свою миссию строго ограничено законодательством.

**Диссертации и авторефераты диссертаций являются авторскими работами** и подпадают под действие закона РФ “Об авторском праве и смежных правах” (1993 г.). **Авторы диссертаций и авторефератов, в соответствии со ст. 16 этого закона, обладают личными имущественными и личными неимущественными правами**, в том числе правом на опубликование и на воспроизведение.

**В договоре между автором и РГБ регламентируются права Библиотеки на безвозмездное предоставление** электронных копий диссертации и автореферата: в открытом доступе **на сайте Библиотеки в сети Интернет**, для доступа читателям Библиотеки **в интранет** или только специальному персоналу Библиотеки в электронном хранилище Библиотеки (архиве). В случае передачи автором в ЭБД электронной версии он гарантирует в тексте договора **идентичность** (понятие определяется в глоссарии к проекту[2]) переданного полного текста диссертации и/или автореферата их печатным версиям, поступившим на хранение в Библиотеку. В случае пересылки автором договора по почте подпись автора заверяется отделом кадров по его месту работы и печатью.

**Получается, что обеспечение полноты базы данных** электронной библиотеки диссертаций РГБ пока (без изменения законодательной базы или порядка присвоения ученых степеней) – невыполнимая задача. Действительно, автор вправе решать, стоит ли помещать его работу в Интернет на всеобщее обозрение или опубликовать иным способом.

На данном этапе разработчики проекта приняли, пожалуй, все реальные меры для поднятия энтузиазма авторов. С диссертантами и представителями советов по защите, доставляющих тома и документацию в РГБ, работают сотрудники отдела диссертаций. Возможные **стимулы для передачи работ в ЭБД** указаны в нашем объявлении для авторов (приобретение работой и автором широкой известности, формирование научных контактов, бессмысленность неправомерного заимствования при возможности для любого эксперта или руководителя в любой момент увидеть диссертацию на нашем сайте). Понятно, что эти стимулы слабы, по сравнению, скажем, с правилами защиты в некоторых зарубежных странах. Например, в некоторых университетах Германии автор предпочтет передать в вуз электронную версию плюс 4 печатных экземпляра работы, чем 20-50 печатных экземпляров. В противном случае, он не сможет получить степень. Опасения наших авторов по поводу плагиата особенно непонятны в связи с тем, что они обязаны до защиты опубликовать основное содержание диссертаций. Автор может отказаться передать диссертацию в ЭБД и по таким простым причинам, как отсутствие у него полной электронной версии работы, неорганизованность, неуверенность в качестве своей работы, намерение издать книгу, занятость, усталость после процесса защиты и обыкновенная лень.

РГБ периодически обсуждает организационные вопросы с Министерством образования и ВАК.

Большая удача – публикация объявления в упомянутом Бюллетене ВАК. Этот факт, по сути, обозначает официальное признание нашей ЭБД, что снимает распространенный вопрос: а как к этому отнесется ВАК?

Важная причина отказа от передачи работ в ЭБД РГБ – существование собственных электронных коллекций диссертаций в некоторых вузах или научных институтах, что пока не особенно широко распространено, но принято нами во внимание для дальнейшего развития проекта. В настоящее время ведутся переговоры с различными вузами о сотрудничестве. Они будут по договорам передавать в нашу электронную библиотеку файлы диссертаций и авторефератов и возьмут на себя правовые отношения с авторами. РГБ предлагает, со своей стороны, передавать результаты обработки файлов или создавать в вузах “зеркальные” базы данных. Если вуз будет настаивать при заключении договора на хранение полных текстов только в своей электронной библиотеке, возможна передача в ЭБД РГБ только адресов файлов на соответствующих сайтах.

### **5 Технология формирования базы данных и доступ к ней**

Полученные Библиотекой от автора электронные версии диссертаций и авторефератов регистрируются и преобразуются в PDF-файлы [8]. Электронная диссертация или реферат снабжается электронным титульным листом, размещается на сайте РГБ в библиотеке OREL. Для электронных документов составляются библиографические записи в ЭК OREL. Структура электронного адреса разработана таким образом, чтобы можно было присваивать его на стадии каталогизации. В ближайшем будущем планируется вводить БЗ электронных диссертаций и авторефератов в общий ЭК РГБ, который функционирует на Веб-сайте Библиотеки [1]. Он представляет собой модуль интегрированной автоматизированной информационно-библиотечной системы РГБ на основе программного обеспечения ALEPH (система, разработанная фирмой EX LIBRIS, Израиль, поддерживает информационную систему РГБ в целом).

Доступ к базе данных полных текстов диссертаций и авторефератов на Web-сайте РГБ [1] обеспечен различными путями: путем поиска с первой страницы, путем выхода в раздел “Электронные ресурсы” в верхнем меню, а затем в левом вертикальном меню – в “Электронную библиотеку диссертаций”; путем прямого выхода в OREL [4] и поиска в этой библиотеке.

**Особое внимание уделяется обеспечению многоаспектного тематического поиска** с учетом иерархических связей между предметами поиска. В последнее время в библиографические записи диссертаций и авторефератов вносятся словесные формулировки индексов ББК, описывающие темы документов, как это с 1998 года делается при обработке книг.

Пример: смысловое содержание работы “Процесс формирования мозга и влияние на него информационных полей окружающей среды” описано в ЭК тремя индексами ББК, для которых в БЗ документа приводятся словесные формулировки. При этом отображается полная цепочка иерархических связей между темами, соответствующая индексу:

- 1) *Психология -- Развитие психики. Генетическая психология -- Развитие мозга и развитие психики в филогенезе*
- 2) *Биологические науки -- Физиология, биофизика и биохимия животных и человека -- Физиология, биофизика и биохимия нервной системы -- Центральная нервная система -- Головной мозг -- Экологическая физиология головного мозга*
- 3) *Здравоохранение. Медицинские науки -- Лечебники.*

Данный документ может быть найден по любому сочетанию перечисленных слов и их частей (*психолог? И экологи?; биофизика И головн? И мозг?; медицин? И развити? И психи?; биологи? И мозг?*)

### **6 Краткое сравнение с зарубежными проектами ЭБД**

Дальнейшее развитие нашего проекта во многом ориентировано на зарубежные тенденции в этой области. Создание электронных библиотек диссертаций и авторефератов представляет собой одно из наиболее активно развивающихся направлений в зарубежной информационной и библиотечной деятельности [8]. В рамках нашего проекта были изучены наиболее серьезные зарубежные проекты, в частности: опыт разработки некоммерческого международного проекта Networked Digital Library of Theses and Dissertations (NDLTD), начинавшегося как корпоративный проект американских университетов [9]; двух университетских проектов во Франции, один из которых становится международным [10]; проекта в Австралии [11]; особенно близкого для нашей задачи проекта “Dissertationen Online” (DissOnline) в Германии, который координируется Die Deutsche Bibliothek [12] (перспективные разработки, связанные с представлением диссертаций в XML и подготовкой DTD, проводятся для данного проекта в Берлинском университете им. Гумбольдта); университетских проектов в Финляндии [13]; коммерческого проекта предоставления в доступ электронных диссертаций компании UMI в ProQuest [14].

**В нашем и зарубежных проектах можно выделить следующие общие моменты:**

- представление полных текстов публикуемых в ЭБД трудов в формате *PDF* и вывод в нем для пользователей;
- тенденция к формированию *XML*-документов, в частности – для долговременного хранения данных;
- ориентация при разработке метаданных (включая DTD) на стандарты системы *NDLTD* и формат *Dublin Core*;

- тенденция к вводу библиографических записей для электронных документов **в общие электронные каталоги** с указанием в библиографических записях электронного адреса;
- **строгое соблюдение авторских прав** путем подготовки договоров с авторами или с вузами, которые поставляют электронные диссертации или сведения о них в ЭБД.

**Различия между нашим проектом и зарубежными сводятся преимущественно к следующему:**

- несовпадение принципов и технологии формирования и хранения баз данных ЭБД в связи с кардинальными различиями в **правилах присвоения ученых степеней**, процессах прохождения диссертаций и организации их депозитарного хранения;
- различия в правилах **оформления** диссертаций;
- различия в **стимулах** для публикации работ в ЭБД;
- разный **уровень компьютеризации** научной деятельности (фактор с наименьшим влиянием).

При всех различиях в практике защиты и хранения диссертаций разработчики зарубежных ЭБД также признают, что **обеспечение полноты баз данных** таких библиотек пока (без изменения законодательной базы) – невыполнимая задача.

## 7 Развитие ЭБД на базе ИСИР

Изучение зарубежных тенденций развития ЭБД позволило разработчикам утвердиться в своих намерениях, сложившихся в самом начале работы над проектом благодаря общению с отечественными специалистами: создать, наряду с простой версией ЭБД, которая описана выше, систему, работающую с электронным хранилищем и расширяющую поисковые возможности электронной библиотеки.

Для преобразования метаданных о диссертациях и авторефератах в формат хранения базы данных был разработан XML-формат описания текстов диссертаций. Кроме функции обмена данными, этот формат преследовал и другую цель: обеспечить возможность представления пользователям основной информации о диссертациях не только в виде pdf-файлов, но и в виде легковесных XML/HTML-страниц.

О преимуществе XML-представлении данных [15] написано много. В частности, важен аспект универсальности этого языка. Имея формализованное с его помощью описание структуры данных системы, можно в дальнейшем наладить процесс обмена информацией практически с любой системой, поддерживающей средства импорта-экспорта данных в формате XML. Немаловажное значение для нас имел и вопрос долговременного хранения данных в текстовом формате.

В рамках настоящего проекта было принято решение использовать XML-разметку только для некоторых компонентов диссертации, составляющих так называемый **экстракт диссертации**. В него включаются: все элементы титульного листа (для

автореферата – обложки и ее оборотной стороны); оглавление, размеченное по всем разделам; список использованной литературы, размеченный поэлементно с выделением авторов, заглавий, источников и всех остальных элементов библиографических записей использованных в диссертации работ (недифференцированно).

При формировании набора элементов формата за основу был взят широко используемый в научном сообществе формат Dublin Core [16], кроме того, учитывалась необходимость установления соответствия элементов данных международному формату представления библиографических данных MARC21, используемому в электронных каталогах РГБ.

Разработанный XML-формат представления данных размещен в настоящее время на сайте РГБ в виде DTD-описания [17].

Что касается поиска, прежде всего хотелось выйти за рамки обычных возможностей электронных каталогов. Кроме того, разработчики посчитали интересным реализовать в ЭБД функцию специальной обработки массивов цитируемой в диссертациях литературы. Поэтому в рамках проекта РГБ-РФФИ подготовлен пилотный проект второй версии ЭБД. Для этой цели был выбран **программный продукт ИСИР РАН (Интегрированная система информационных ресурсов Российской академии наук)** [18] в качестве будущей основы для реализации программного обеспечения ЭБД, позволяющего поддерживать базу данных, умеющую работать с электронными XML – документами, и обеспечивающего разработанные в рамках проекта алгоритмы поиска информации. По договору с Вычислительным центром им. А.А. Дородницына РАН на выполнение научно-исследовательской работы программный продукт частично адаптирован и тестировался на примерах (около 15 текстов).

Таким образом, для формирования электронного хранилища на основе программного обеспечения ИСИР РАН предоставляются:

- электронные версии диссертаций и авторефератов, получаемые от авторов, с электронными титульными листами в формате **PDF**;
- экстракты работ в виде **XML**-файлов, включающие метаданные.

Как указано выше, разработчики проекта приняли решение сделать основным форматом для вывода полных текстов диссертаций формат **PDF**. Для получения данных, необходимых для загрузки информации в ИСИР, в системе будет обеспечено конвертирование метаданных, представленных в текущей реализации ЭК в международном формате **MARC21**, в записи на языке разметки **XML**.

Для обеспечения целостности и согласованности информации разрабатываемой системы и Электронного каталога(ЭК) РГБ библиографическая запись для электронного документа, хранящаяся в ЭК, будет содержать ссылку как на полный PDF -текст диссертации (где бы ни хранился этот файл – на

сайте РГБ или на сайте ЭБД-ИСИР), так и на ее XML-представление.

Итак, перед разработчиками программного продукта была поставлена задача разработать *систему, отвечающую следующим требованиям*:

- система предоставляет пользователям удаленный доступ через Интернет и разработана на основе современных порталных технологий;
- система обеспечивает единый пользовательский интерфейс для доступа к соответствующим, возможно, гетерогенным сервисам и источникам информации:  
с точки зрения пользователя система обеспечивает единство средств доступа и обслуживания; пользователю должна предоставляться согласованная система просмотра, поиска, обнаружения ресурса, его заказа, доставки;
- система имеет средства для аутентификации и пользователей как индивидуумов, так и групп, средства контроля прав доступа на соответствующих уровнях. Предоставляется определенный ограниченный доступ для неавторизованных пользователей;
- система поддерживает и использует достаточно детализированные профили пользователей, основанные на соответствующих открытых стандартах;
- система имеет средства рубрикации ресурсов, обеспечивая их тематическую организацию;
- система предусматривает средства ведения статистической и административной информации, например, относительно объемов и типов транзакций;
- система обеспечивает относительно простое управление, не требующее привлечения экспертов для ее поддержки: системные администраторы должны иметь возможность добавлять и подключать дополнительные ресурсы;
- система использует современные открытые стандарты работы с данными;
- система предоставляет средства пакетной загрузки и выгрузки данных.

За основу для разработки такой системы была взята технология Интегрированной системы информационных ресурсов (ИСИР) [19]. Этот выбор был обусловлен следующими *возможностями ИСИР*:

- в ИСИР разработана детализированная структура данных, охватывающая основные типы информационной системы типа “цифровая библиотека” (публикация, персона, организация и т.д.) и связи между ними;
- интерфейсы ИСИР легко модифицируются для нужд конкретной системы;
- в ИСИР существуют мощные средства поиска, манипуляции данными, авторизации, разграничения доступа и т.д.;
- структура данных ИСИР легко модифицируется и расширяется согласно специфическим требованиям новой области данных;

- в ИСИР используются современные технологии (портальные решения, XML, XSLT, RDFS, JAVA);
- в ИСИР разработаны средства импорта-экспорта данных;
- ИСИР позволяет развернуть систему на разных платформах в зависимости от требований заказчика;
- имеется опыт успешной эксплуатации системы.

*Схема данных и стандартные интерфейсы ИСИР* были *модифицированы* в соответствии с требованиями заказчика: изменен атрибутивный (сложный) поиск, добавлен поиск слов и выделен поиск цитируемой литературы; изменены форматы выдачи данных. Кроме того, был модифицирован загрузчик для экспорта данных в систему в разработанном XML-формате.

*Структура данных ИСИР* была расширена и модифицирована в соответствии с требованиями предметной области:

- расширено понятие типа ресурсов “публикация”;
- введены новые типы связей между ресурсами “публикация”, “персона” и “организация” (Рисунок 1).



**Рисунок 1 Ресурсы ИСИР-ЭБД.**

Остановимся подробнее на внесенных модификациях. В рамках технологии ИСИР ресурс “публикация” описывает основные элементы, характерные практически для любого типа публикации. Задача достаточно подробного описания специфических типов публикации – диссертации и автореферата – потребовала введения дополнительных свойств, описывающих публикацию.

Для спецификации ресурса “публикация” ИСИР-ЭБД были введены следующие *дополнительные свойства*:

- **Типизированные ключевые слова.**  
В схеме данных присутствует несколько типов ключевых слов:  
ключевые слова в стандартном понимании этот термина;  
ключевые слова, описывающие персону, именуемую отношение к диссертации;

ключевые слова, описывающие организацию, имеющую отношение к диссертации; ключевые слова, описывающие временные периоды, имеющие отношения к диссертации; ключевые слова, описывающие географические понятия, имеющие отношение к диссертации. В зависимости от функциональности клиента, эти типы могут различаться либо рассматриваться просто как общие ключевые слова.

- **Содержание (оглавление).**

Это текстовый атрибут, содержащий оглавление диссертации. Наличие такой информации позволяет пользователю точнее определить характер и релевантность найденных ресурсов.

- **Список литературы.**

В соответствии с парадигмой ИСИР публикации, указанные в списке литературы, можно рассматривать как самостоятельные ресурсы типа “публикация”. Такой ресурс обладает достаточно богатым описанием: приводится его название, авторы, год публикации, издательство, для статьи в издании - страницы. При соответствующем представлении входных данных (подробной разметке информации) можно создать в системе все необходимые ресурсы для заведения новой публикации, указанной в ссылке – или идентифицировать данные с уже имеющимися в системе ресурсами (персонами-авторами, организацией-издательством, публикацией), что предоставляет пользователю удобный усовершенствованный интерфейс для работы с данными. Но такая работа требует, во-первых, участия человека (оператора) для проверки правильности установленных связей (например, сложно в автоматическом режиме идентифицировать персону-автора по фамилии и инициалам) и, во-вторых, очень подробно и правильно размеченной секции “Список литературы”. Поэтому при модификации схемы данных было принято следующее решение: предоставить возможность хранить как ссылки на цитируемые публикации, так и текстовый атрибут, содержащий строку со ссылкой на публикацию в том виде, в котором она дана в исходном тексте. Таким образом, в системе можно поддержать и упрощенную схему хранения данных о ссылках на публикации. Связи между диссертацией и публикациями из списка литературы хранятся в отсортированном виде и выдаются пользователю в том же порядке, как и в печатном источнике.

- **Классификаторы и справочники.**

В отличие от других типов публикаций, диссертации и авторефераты рубрицируются классификатором ВАК. В ИСИР этот классификатор поддерживается, и для диссертаций были установлены связи между публикациями и рубрикатом ВАК. Дополнительным справочником для публикаций стал справочник ученых степеней – предмет защиты диссертации.

- **Новые связи между ресурсами.**

Для связи публикации и персоны были введены следующие типы отношений: научный руководитель, оппонент, рецензент; публикации и организации: учреждение, где защищалась диссертация (ученый совет), ведущая организация и учреждение, где выполнена диссертация.

## 8 Заключение

В заключение хочется отметить несколько важных вопросов, связанных с переходом на новую версию ЭБД. Для успешной реализации новой версии ЭБД мы считаем целесообразным в дальнейшем преобразовать ее в распределенную электронную библиотеку полных текстов диссертационных работ с обязательным хранением метаданных и адресов текстов в центральной базе данных РГБ.

Наиболее сложной из поставленных в организационном плане задач можно считать предполагаемое внедрение специальных html-форм для ввода в них самим автором текстов диссертаций, авторефератов и сведений о себе. Разработка и внедрение механизма создания единообразного формализованного представления диссертационных работ необходимо как для обеспечения автоматической генерации XML-представлений диссертаций, так и для формирования других типов документов (например, Word-документов). Примером может служить форма, разработанная в Берлинском университете им. Гумбольдта, а также некоторые отечественные разработки для других типов текстов. Наличие такого сервиса позволит повысить качество поступающей в ЭБД информации, а также вовлечь авторов диссертаций в процесс активного взаимодействия с ЭБД.

## Литература

- [1] Сайт РГБ.  
<http://www.rsl.ru>
- [2] Web-страница проекта.  
<http://rsl.ru/pub.asp?10.htm>
- [3] Сайт ЮНЕСКО - UNESCO Clearing House on Electronic Theses and Dissertations (ETD).  
<http://www.eduserver.de/unesco/>  
<http://dbs.bbf.dipf.de/db/institutions.html>
- [4] Сайт OREL  
<http://orel.rsl.ru/index.shtml>
- [5] Груздев И.А., Перли Б.С., Лавренова О.А. Электронная библиотека РГБ – составная часть РГБ // Электронные библиотеки: перспективные методы и технологии, электронные коллекции. Сб. докладов. Третья Всероссийская конференция RCDL'2001, Петрозаводск, 11-13 сентября 2001 г. – Карельский центр РАН, 2001. – С.232-236
- [6] Лавренова О.А. Проект электронной библиотеки диссертаций в РГБ // Электронные изображения и визуальные искусства: Культурное наследие для всех: новые технологии в музеях, галереях, библиотеках, архивах. Материалы кон-



- ференции *EVA'2001* Москва (3-8 декабря 2001 г.). – Москва, 2001. – С. 13-2-1 – 13-2-3
- [7] Лавренова О.А. Электронная библиотека диссертаций в РГБ // Материалы 9-й Международной конференции “Крым – 2002” “Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества”. – Су-дак; Ялта. 2002. – Т.1 - С. 235-239
- [8] Лавренова О.А. Электронная библиотека диссертаций и международный опыт.  
<http://www.conf.cultivate.ru/2003/russian/spkdok.html>  
[http://www.conf.cultivate.ru/2003/russian/dok\\_107.html](http://www.conf.cultivate.ru/2003/russian/dok_107.html)
- [9] Networked Digital Library of Theses and Dissertations (NDLTD).  
<http://www.ndltd.org>  
<http://www.theses.org>
- [10] Systeme universitaire de documentation.  
<http://www.sudoc.abes.fr>  
University of Marne-la-vallée (UMLV).  
<http://www.univ-mlv.fr>
- [11] Australian Digital Theses Program (ADT).  
<http://adt.caul.edu.au/>
- [12] Project "Dissertation Online".  
<http://www.dissonline.de>
- [13] Helsinki University of Technology Library.  
<http://lib.hut.fi/Diss/>
- [14] E-thesis: Electronic Publications at University of Helsinki.  
<http://ethesis.helsinki.fi/english.html>
- [15] Extensible Markup Language (XML) 1.0, W3C Recommendation, 10 February 1998.  
<http://www.w3.org/TR/1998/REC-xml-19980210.html>
- [16] The Dublin Core Metadata Initiative (DCMI)  
<http://www.dublincore.org>
- [17] DTD-описание формата метаданных.  
[http://rsl.ru/pub.asp?10\\_12.htm](http://rsl.ru/pub.asp?10_12.htm)
- [18] Кулагин М. В. , Серебряков В. А. , Бездушный А. Н. , Галочкин М. П. , Меденников А. М. , Агошков С. В. Интегрированная Система Информационных Ресурсов (ИСИР) РАН - подход к созданию интегрированных цифровых библиотек// Международная научная конференция "Электронные библиотеки: перспективные методы и технологии, электронные коллекции", Санкт-Петербург, 1999.
- [19] Кулагин М. В. , Серебряков В. А. , Бездушный А. Н. , Галочкин М. П. , Меденников А. М. , Агошков С. В. Интегрированная Система Информационных Ресурсов (ИСИР) РАН - подход к созданию интегрированных цифровых библиотек// Международная научная конференция "Электронные библиотеки: перспективные методы и технологии, электронные коллекции", Санкт-Петербург, 1999.
- [20] Бездушный А.Н., Власова С. А. , Каленов Н. Е. , Меденников А. М. , Серебряков В. А. Подход к интеграции информационных коллекций в ИСИР РАН// Международная научная

конференция "Электронные библиотеки: перспективные методы и технологии, электронные коллекции", Протвино, 26-28 сентября, 2000.

### **Structure and Implementation of the RSL Dissertations Digital Library.**

Lavrenova O.A., Vezhnevets A.A.

The article describes the digital library of Electronic Theses and Dissertations (ETDs) project supported by the Russian Foundation for Basic Research and developed by RSL. The following aspects are discussed: the main results; common matters in the Russian and foreign projects and differences between them; the copyright questions, the assurance of the completeness of the data base and possible incentives; the data processing; access to the data base of full texts; the assurance of the subject search; further development of the project on the base of the Integrated System of Information Resources of the Russian Academy of Sciences. The main result is that the Library has begun to collect dissertations and theses in digital form from authors who defended them before specialized scientific councils in this country, and who wish to give their works for the digital library of the RSL. The Library concludes an agreement with each author by which an open access to a dissertation and /or theses text is provided through the RSL Web-site in Internet or Library intranet – depending on the terms of the agreement with each author. The digital versions of dissertations and theses received from the authors are transformed into the PDF files. The functioning of the digital dissertations library is maintained by the following components of the on-line public access catalogue (OPAC): authority files, in particular data on persons (authors, executives); data on corporate bodies (institutions); the machine-readable national library - bibliographic classification (BBK). The digital library of ETDs on the base of the ISIR software (pilot project) is based on fulltext database of theses and dissertations in PDF and excerpts (extracts) of works in the form of XML files.

---

\* Проект электронной библиотеки диссертаций и авторефератов (ЭБД) Российская государственная библиотека (РГБ) разрабатывает при поддержке Российского фонда фундаментальных исследований (РФФИ), проект 01-07-90310