Концептуальная модель электронной библиотеки

© Резниченко В.А., Проскудина Г.Ю., Кудим К.А.
Институт программных систем НАН Украины, г. Киев 03187, пр. Академика Глушкова, 40
reznich@isofts.kiev.ua, gupros@isofts.kiev.ua, kuzma@isofts.kiev.ua

Аннотация

Работа посвящена задаче создания концептуальной модели электронной библиотеки. Обсуждаются некоторые известные связанные проекты – CIDOC CRM, FRBR, DELOS DLRM. Предложен оригинальный вариант информационной составляющей концептуальной модели.

"Пора подумать, – Морж сказал, –
О множестве вещей:"

Л.Кэрролл "Алиса в Зазеркалье"

1 Введение

Появление новых электронных библиотек (ЭБ), увеличение числа хранимых в них документов и повышение качества предоставляемых ими услуг способствует развитию науки, облегчая, а иногда и просто открывая единственно возможный доступ к источникам информации для ученого, предоставляя ему замечательное средство донести плоды своей деятельности до широчайшей аудитории. В последние несколько лет при нашем непосредственном участии научное сообщество Украины продвинулось в этом направлении. В частности, в прошлом году создан портал периодических изданий НАН Украины NASPLIB[1]. Два года назад создана ЭБ Института программных систем НАН Украины ISS EPrints[2]. В первом случае использовалось программное обеспечение DSpace, во втором – EPrints. Обе системы были полностью украинизированы. Были отработаны основные сценарии использования, создан ряд методик и рекомендаций по созданию и использованию электронных библиотек на основе данных программных систем. Были также изучены программные продукты Greenstone[3] и Fedora[4]. Этот опыт оказался очень ценным для понимания современного состояния дел в мире программных систем ЭБ.

В настоящее время нет какой-либо универсальной ЭБ, которая отвечала бы всем требованиям и ожиданиям пользователей. Анализ существующих систем ЭБ [1-3] 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 гетерогенности ЭБ и игнорирования нужд их 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.

Решить эти и другие возникающие проблемы на первом этапе поможет аккуратное и полное рассмотрение области ЭБ. Именно для этого создаются концептуальные модели, обобщающие накопленный опыт в сфере создания и использования ЭБ.

В последнее время в мире предпринимаются усилия по полному и всестороннему описанию сферы ЭБ. Во втором разделе мы обсуждаем следующие известные модели и стандарты, которые могут применяться для описания ЭБ в целом и ее частей: CIDOC CRM [4], FRBR [5], DELOS DLRM [6].

Третий раздел посвящен описанию информационной составляющей разрабатываемой нами концептуальной модели ЭБ. Мы рассматриваем ЭБ как информационную систему, поэтому в дальнейшем планируется дополнить описание модели пользовательской и функциональной составляющими, которые не затронуты в данной работе.

2 Обзор известных проектов

Нацелившись на описание информационного пространства библиотечной системы, мы естественно не могли не рассмотреть что-то подобное, что уже сделано или делается сегодня в мире.

2.1 CIDOC CRM

Концептуальная эталонная модель (Conceptual Reference Model, CRM) CIDOC, разработанная Международным комитетом по документации Международного совета музеев (The International Committee

Труды 11[ой] Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» - RCDL'2009, Петрозаводск, Россия, 2009.

for Documentation of the International Council of Museums, ICOM-CIDOC), предназначена для интеграции, посредничества и обмена информацией в области мирового культурного наследия и связанных областей.

Представляя определения и формальную структуру описания неявных (implicit) и явных (explicit) сущностей и отношений, модель претендует на общий язык для экспертов и разработчиков, формулирующих требования к информационным системам, способствует общему пониманию информации, обеспечивает интеграцию, посредничество и обмен информацией между музеями, библиотеками, архивами [7, 8].

До 1994 года разрабатывалась ER-модель для музейной информации, начиная с 1996 года подход разработки модели сместился к методологиям объектно-ориентированного моделирования и привел в 1999 году к появлению первой Концептуальной эталонной модели CIDOC. С 2000 года начался процесс стандартизации, который успешно завершился принятием стандарта ISO 21127:2006 – "Эталонная онтология для обмена информацией культурного наследия" (A reference ontology for the interchange of cultural heritage information).

Разработчики CIDOC CRM поставили своей целью написать стандарт модели, пригодной как для машинной обработки, так и для легкого понимания человеком. Модель совместима с формализмом RDF.

Версия 4.2.4 модели CIDOC CRM [4] состоит из 87 классов и 148 свойств, описывающих предметы, понятия, людей, события, место, время и их отношения. CIDOC CRM предлагает только высокоуровневые понятия, описывающие сущности и их связи, и никак не связана с документированием либо реализацией таких систем.

На рис. 1 представлена часть иерархии классов CIDOC CRM. Все классы, за исключением класса *Простое значение* (*Primitive Value*) и его подклассов, прямо или опосредовано являются подклассами класса *E1 Сущность CRM*, охватывающего все сущности, которые могут быть описаны в CIDOC CRM.

Модель может быть расширена добавлением необходимых для конкретной задачи сущностей в иерархию классов.

Резюмируя рассмотрение данной модели, отметим, что для наших целей это – нужный и полезный стандарт. Важным преимуществом стандарта является его формальный подход. Он вполне может служить основой для информационной составляющей концептуальной модели ЭБ. Конечно, стандарт нуждается в расширении более конкретными сущностями, которые часто используются во многих ЭБ. Кроме того, CIDOC CRM не охватывает пользовательского и функционального аспектов ЭБ.

2.2 FRBR и FRBRoo

Независимо от CIDOC CRM в 1991-1997 годах Международной федерацией библиотечных ассоциаций и учреждений (International Federation of Library Associations and Institutions, IFLA) была разработана ER-модель "Функциональные требования к библиографическим записям" (Functional Requirements for Bibliographic Records, FRBR) как обобщенное представление библиографического универсума, независимого от какого-либо кода каталогизации или реализации.

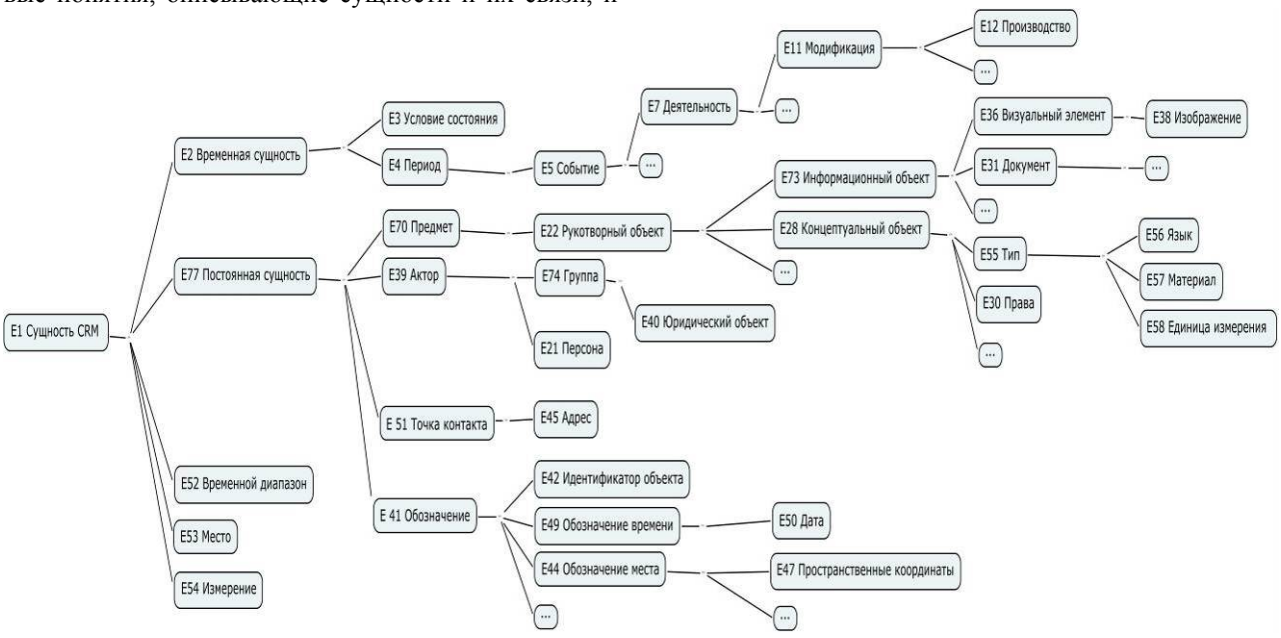

Рис. 1. Часть иерархии классов в модели CIDOC CRM

В 1998 году модель была опубликована [5]. В настоящее время IFLA продолжает контролировать приложения модели FRBR и поддерживает ее использование и развитие.

FRBR включает описание концептуальной модели (сущности, их отношения и атрибуты), предлагает универсальные библиографические записи для всех типов материалов и пользовательских задач, связанных с библиографическими ресурсами, описанными в каталогах, библиографиях и других библиографических инструментах [9, 10].

Модель FRBR различает три группы сущностей (рис. 2):

- для описываемых объектов: *произведение (work)*, *выражение (expression)*, *воплощение (manifestation)*, *экземпляр (item)*;
- для описателей-субъектов: *человек (person)* и *организация (corporate body)*;
- для описателей-объектов: *концепт, объект, событие* и *место (concept, object, event, place)*.

Ниже приведен пример [11] экземпляров сущностей *произведения* (w1) и его *выражений* (e1-e2):

w1 *Tennis--bis zum Turnierspieler* Элльвангера

- e1 оригинальный текст на немецком языке
- e2 перевод на английский язык Венди Джилл
- ...

Большое внимание в модели уделено отношениям между сущностями.

Отношения могут быть отражены в библиографических записях многими способами. Те, что изображены на ER-диаграмме FRBR (рис. 2), описывают логические связи между сущностями и часто реализуются простой конкатенацией одной сущности с атрибутами связанной сущности в одной записи.

Помимо логических связей в модели выделена группа так называемых *контентных* связей (для первой группы сущностей). Они идентифицируют основные типы отношений, которые существуют между экземплярами сущности одного типа (например, сущности *произведения*) или между экземплярами разных типов сущностей (например, экземпляров сущностей *произведение* и *воплощение*). Например, в группе отношений *произведение-произведение* выделены такие типы отношений: *имеет адаптацию* (свободный перевод); *имеет приложение* (сходство, соответствие), *имеет продолжение*; *имеет резюме* (обзор, аннотацию); *имеет преобразование (*стихотворную форму*)*; *имеет имитацию* (пародию). В группе отношений *выражение-выражение* перечислены следующие типы отношений: *имеет сокращение* (корректировку, уплотнение); *имеет пересмо*тр (исправленную редакцию, расширенную редакцию); *имеет перевод* (буквальный перевод) и некоторые другие типы отношений, касающиеся музыкальных произведений.

И наконец, отношения *часть/целое* и *часть в части* также представлены в модели FRBR.

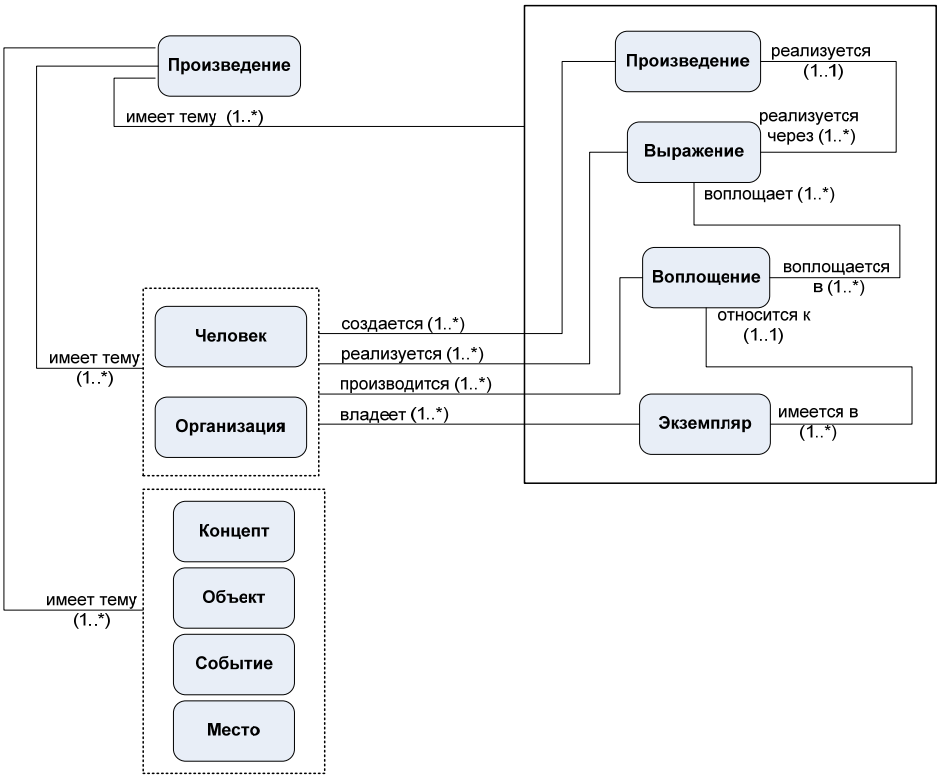

Рис. 2. Модель FRBR

В 2006 году был опубликован первый полный проект модели FRBRoo, т.е. объектно-ориентированной версии FRBR, согласованной с CIDOC CRM. Эта формальная модель предназначена для представления основной семантики библиографической информации, интеграции и обмена библиографической и музейной информацией.

Главное новшество FRBRoo – реалистичная, явная модель процесса интеллектуального творчества, которая еще должна получить свое дальнейшее развитие для библиотекарей и ученых [12].

Следует отметить, что в FRBR границы между различными типами основных сущностей (произведение и выражение) размыты и окончательное решение по тому, к какому типу отнести тот или иной объект, отдается на откуп каталогизатору. Кроме того, сущностей этих совсем немного и явно недостаточно для большинства конкретных библиотечных приложений. Для нас основной интерес представляет очень богатый набор атрибутов и отношений в этой модели. Как и в случае CIDOC CRM, в соответствии с решаемыми задачами, FRBR применима только для описания информационной составляющей концептуальной модели.

2.3 DELOS DLRM

Группа специалистов ассоциации в сфере ЭБ DELOS в 2006-7 гг., основываясь на анализе имеющихся библиотечных систем [3], где большое внимание было уделено функциональным возможностям современных ЭБ, начали разработку эталонной модели ЭБ (Digital Library Reference Model, DLRM) [6]. Цель проекта – разобраться с фундаментальными понятиями, существенными объектами и их отношениями, стандартными функциональными и структурными блоками и процессами, из которых состоит универсум ЭБ. Эталонная модель предназначена для разработки более узких моделей с конкретной архитектурой для последующей реализации программных систем.

Прежде всего, в модели было выделено три понятия для разграничения того, что обычно называется ЭБ:

- *ЭБ* – конкретная ЭБ с ее пользователями, правилами, содержимым, интернет-сайтом и ведущей организацией. Например: библиотека института программных систем ISS EPrints http://eprints.isofts.kiev.ua;
- *система ЭБ* – программное обеспечение, на основе которого создаются ЭБ. Например: EPrints 3.0.
- *система управления ЭБ* – программное обеспечение для создания и управления системами ЭБ. Например: система OpenDLib[5].

Далее модель DELOS DLRM рассматривается в ролевом аспекте, т.е. с точки зрения разных категорий пользователей:

- конечный пользователь ЭБ;
- разработчик ЭБ;
- системный администратор ЭБ;
- разработчик приложений для ЭБ.

Соответственно DELOS DLRM имеет четыре уровня пользовательских представлений.

Весь универсум ЭБ разбит на шесть высокоуровневых ключевых областей (рис. 3):

- контент;
- пользователь;
- функциональные возможности;
- качество;
- политики;
- архитектура;

и несколько дополнительных. Эти шесть областей объединены в одну область ресурса. В каждой из них вводятся и определяются свои сущности и их свойства.

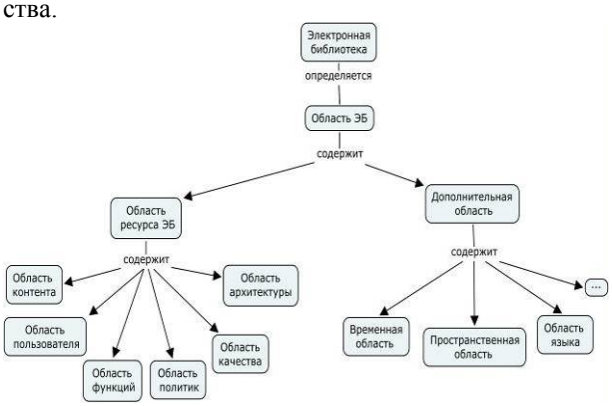

Рис. 3. Иерархия областей ЭБ в модели DELOS DLRM

Теперь вкратце рассмотрим наиболее важные области ЭБ и их структуру.

Область ресурса ЭБ – наиболее общая область в данной модели, представляет все сущности и связи, населяющие универсум ЭБ. *Ресурс* – наиболее общее понятие, включающее любую сущность ЭБ. По аналогии с ресурсом в Веб, ресурс – это все то, что может быть идентифицировано, названо или адресовано. Представленная здесь модель ресурса исходит из веб-архитектуры, но дополнена некоторыми аспектами, специфичными для предметной области ЭБ.

Ресурс – абстрактное понятие, в том смысле, что непосредсвенно не имеет экземпляров, он только выражен экземплярами одной из своих специализаций. В частности, экземплярами понятия ресурс в универсуме ЭБ являются экземпляры *информационного объекта* любого типа (например, документы, изображения, видео, мультимедийные объекты, наборы метаданных и аннотаций, потоки, базы даных, коллекции, запросы и результаты запросов), *акторы* (как одушевленные так и неодушевленные сущности), *функции*, *политики*, *параметры качества* ЭБ и *архитектурные компоненты*. Каждый из этих экземпляров представляет главное понятие в своей области, т.о. в представленной модели ЭБ каждая область состоит из ресурсов, а ресурсы – строительные блоки всех областей ЭБ. Каждый ресурс:

- имеет идентификатор;

– организован в соответствии с *форматом ресурса*. Формат здесь выражен *онтологией*. Ресурс может быть сложным и структурированным, поскольку он, в свою очередь, может состоять из меньших ресурсов и иметь связи с другими ресурсами;
– может характеризоваться параметрами качества;
– может регулироваться политиками, управляющими его жизненным циклом;
– выражается через информационный объект;
– может быть описан или дополнен информационным объектом, обычно – метаданными и аннотациями.

С организационной точки зрения, ресурсы могут группироваться в наборы ресурсов, которые рассматриваются как единая сущность. Например, *коллекции* в области контента или *группы* в области пользователя.

Область контента (рис. 4) представляет все объекты, связанные с информацией, которой управляет ЭБ. *Информационный объект* – наиболее общее понятие в этой области, представляет произвольную единицу информации, управляемую в универсуме ЭБ. В DELOS DLRM различают *информационный объект по уровню абстракции,* где заимствуются типы объектов из модели FRBR (произведение, выражение, воплощение) и *информационный объект по связи* – "абстрактный концептуальный контейнер для классов, которые порождают эти объекты", а именно:

– *первичный информационный объект* – информационный объект, который используется самостоятельно, например, текстовые документы, изображения;
– объект *метаданные* – информационный объект, главная цель которого состоит в том, чтобы дать информацию о целевом ресурсе (как правило о первичном информационном объекте);
– объект *аннотация* – информационный объект, главная цель которого состоит в том, чтобы аннотировать целевой ресурс или его часть (на рис. 4 части ресурса соответствует сущность *регион*). Примеры таких аннотаций включают примечания, структурированные комментарии и связи. Объекты аннотации помогают интерпретировать целевой ресурс, содержат либо поддержку, либо возражения, либо более детальные объяснения.

Поскольку информационный объект является ресурсом, то он наследует все вышеперечисленные свойства ресурса.

Информационные объекты также могут быть сложными объектами и могут быть сгруппированы в *коллекции* информационных объектов. Коллекции, в свою очередь, тоже являются информационными объектами, они наследуют все аспекты моделирования информационных объектов и средства их обслуживания, например они могут аннотироваться. Кроме того, коллекция – специализация понятия *набора ресурсов*. Коллекции определяются *критерием отбора* (hasIntension) либо *перечислением элементов* (hasExtension). Другая специализация понятия набор ресурсов в данной модели – *результирующий набор*. В традиционных ЭБ он представляет собой набор документов, которые извлекаются в ответ на *запрос*.

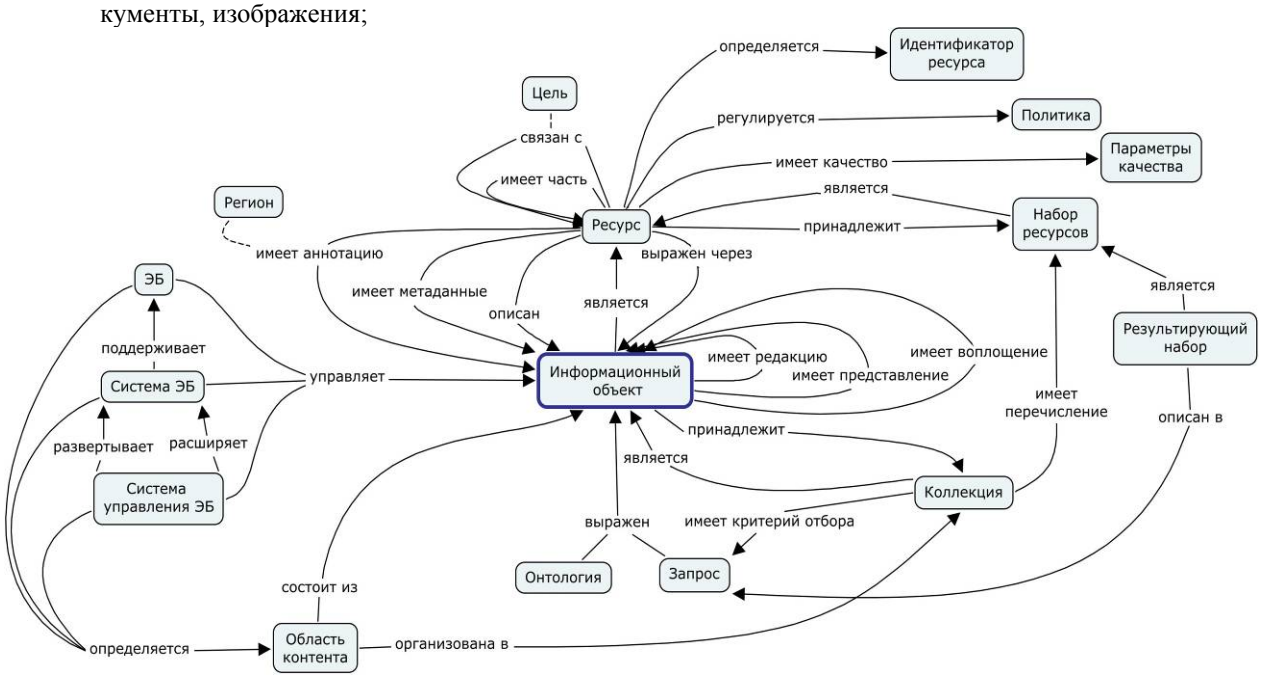

Рис. 4. Область контента ЭБ в модели DELOS DLRM

Область пользователя в DELOS DLRM содержит все объекты, которые являются "внешними по отношению к системе ЭБ и с ней взаимодействуют:

люди и неодушевленные объекты, например, программы или физические инструменты… или даже другая ЭБ может быть среди пользователей ЭБ".

Поскольку главная сущность в этой области – *актор* является ресурсом и следовательно, наследует все его свойства, а именно:

- имеет уникальный идентификатор (идентификатор пользователя);
- организован в соответствие с форматом (модель пользователя);
- благодаря свойствам ресурса композиции и соединения может быть составлен в различные сложные и структурированные группы например, сотрудничество пользователей или соавторов;
- описан или дополнен метаданными и аннотациями.

Область функций представляет наиболее объемную и наиболее открытую часть модели DELOS DLRM, поскольку охватывает всю обработку ресурсов, а также действия пользователей в ЭБ. Здесь наиболее общим понятием является сущность *функция*.

Функция – специфическая задача обработки, которая может быть реализована на наборе ресурсов или одном ресурсе в результате действий отдельного пользователя. Описание функций основано на пользовательском аспекте и ресурсе, представляющем все объекты, вовлеченные в ЭБ. Хотя функции в традиционных моделях ЭБ обычно связываются с контентом в ЭБ и выполняются людьми, здесь, в данной модели, функции могут выполняться неодушевленными пользователями на любом типе ресурсов.

В данной модели ЭБ каждая функция также является ресурсом и потому наследует все его характеристики.

Функции разделены на пять классов:

- доступа к ресурсам;
- управления ресурсами;
- совместной работы;
- управления ЭБ;
- настройки ЭБ.

Подводя итог, нужно признать, что именно модель DELOS DLRM вдохновила изначально нашу работу. Внимательное изучение данной модели помогло не только обозреть всю сферу ЭБ, но и найти некоторые пробелы в самой модели. Вот некоторые из них:

- недостаточно формализованные определения, оставляющие размытыми границы многих сущностей (например, сущности, заимствованные из FRBR, или граница между *метаданными* и *аннотацией*);
- в некоторых местах остаются не ясными критерии выделения сущностей (в частности, область качества наименее убедительна в этом отношении);
- неоднородность описания различных областей ЭБ, скрытая за внешне однообразным описанием (достаточно сравнить простую иерархию области функций со сложной структурой области контента).

К преимуществам DELOS DLRM следует отнести наибольшую полноту охвата среди существующих концептуальных моделей ЭБ.

3 Информационная модель ЭБ

Концептуальная модель должна описывать то, какие сущности могут существовать в данной предметной области (для нас – области электронных библиотек, а точнее – электронных научных библиотек), т.е. существуют в данный момент, существовали ранее или когда-либо смогут существовать. А также она должна фиксировать их правила, связи, что в частности предполагает классификацию сущностей, абстрагирование, обобщение.

Основываясь на рассмотренных выше моделях, учитывая их достоинства и недостатки, мы попытались построить свою модель ЭБ, начав описание с ее информационной составляющей.

3.1 Сущности

На рис. 5 изображена иерархия сущностей или объектов, представленных в электронной научной библиотеке.

Физический объект – корневой объект в представляемой модели, он охватывает все объекты, информация о которых хранится в электронной библиотеке.

Физический объект, как и все другие объекты, обладает *атрибутами*. Набор атрибутов объекта зависит от его типа. Так физический объект имеет следующие атрибуты:

- идентификатор физического объекта;
- название;
- тема;
- ключевые слова;
- версия;
- аннотация.

Эти атрибуты наследуются всеми другими объектами представленной иерархии (рис. 5).

Как правило, в системах ЭБ предусматривается хранение *рукотворных объектов* – основного типа объектов информационного контента, а также некоторых других объектов, имеющих к ним отношение:

- *организации, отделы* организаций и *издательства*, где создавались или публиковались рукотворные объекты;
- люди (на схеме это сущность *человек*), работающие в этих организациях (отделах) – авторы рукотворных объектов;
- *проекты* в рамках которых создаются рукотворные объекты;
- научные *журналы* (периодические издания) и *конференции* их публикующие.

Объект *коллекция* может быть применим к любой совокупности (группировке, агрегации) физических объектов. Физические объекты здесь могут быть лю-

бого типа, т.е. коллекциями могут быть как совокупности физических объектов, так и рукотворных объектов, совокупности организаций, журналов и т.д. Критерии для таких совокупностей могут определяться, например, общностью местоположения, общностью авторов, хронологией, тематикой, происхождением или принадлежностью и т.д. [1]. Коллекции могут содержать любое число объектов и критерии отбора этих объектов со временем могут изменяться.

Организации, как правило, представляют научно-исследовательские институты или образовательные организации. Помимо наследуемых атрибутов этот класс имеет также следующие атрибуты и связи:

- тип организации;
- дата основания;
- местонахождение (страна, город);
- вышестоящая организация;
- руководитель;
- подразделение;
- адрес (почтовый, юридический, сайт, e-mail, телефон).

Класс *человек,* наряду с наследуемыми, имеет также и собственные свойства:

- место работы (организация, отдел);
- пол;
- дата рождения;
- место рождения (страна, регион, город);
- адрес (почтовый домашний, личного сайта, e-mail, телефон);
- ученое звание, ученая степень;
- специальность ВАК (Высшая аттестационная комиссия);
- соавторство.

Приведем также перечень атрибутов для класса *проект*:

- название программы;
- название конкурса;
- период выполнения;
- организация (где выполняется проект);
- руководитель;
- спонсор;
- бюджет.

Журнал и *конференция* – объекты, связанные с публикацией (а значит и с производством) главного вида научной продукции – статьи, одного из представителей класса рукотворных объектов.

Перечень возможных атрибутов для класса *журнал*:

- ISSN (Международный стандартный серийный номер);
- издатель,

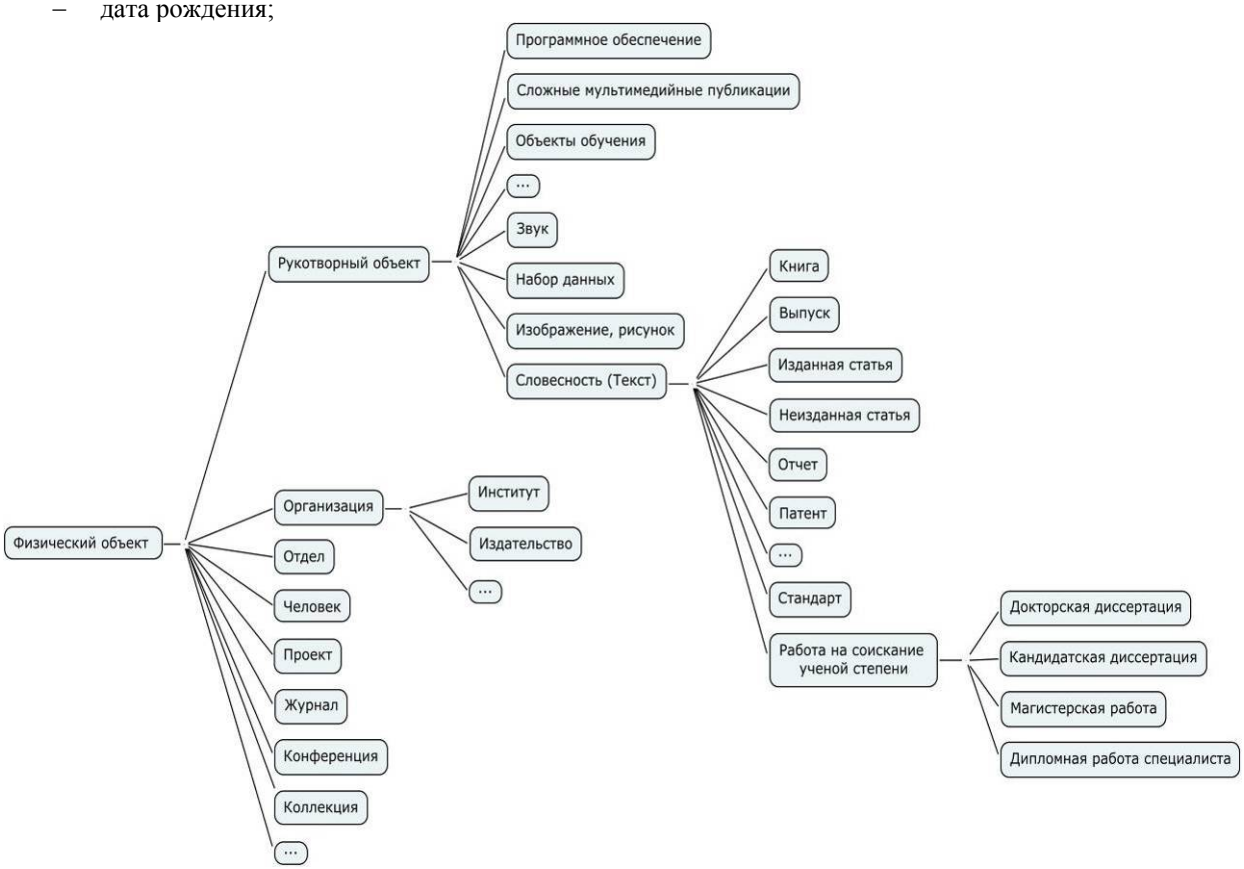

Рис. 5. Иерархия классов в информационной модели научной электронной библиотеки

а также для класса *конференция*:

- дата и место проведения;
- программный комитет;
- ответственная организация.

Рукотворный объект – класс существующих в библиотеке объектов, созданных в процессе научной деятельности людей. К свойствам физического объекта здесь добавляется новое отношение *имеет автора*.

Рукотворные объекты, в свою очередь, подразделяются на:

- текстовые объекты или *словесность;*
- графические объекты (*изображение, рисунок*);
- аудио объекты или *звук;*
- *сложные мультимедийные публикации*;
- *объекты обучения*;
- *наборы данны;*
- *программное обеспечение*.

Объекты *словесность* (текстовые объекты) в представляемой модели используется для обозначения любого электронного текстового контента различных типов – *книги; выпуски* научных периодических журналов; опубликованные в них статьи (*изданная статья*); еще *неизданные* публикации; различные научные *отчеты*; документы, вошедшие в разряд принятых *стандартов*; *патенты*, а также работы, представленные на соискание ученой степени (докторская, кандидатская диссертация, магистерская и дипломная работа). На схеме (рис. 5) для целью простоты и наглядности не показаны такие типы объектов словесности как инструкции, методические материалы, тексты и презентации докладов и выступлений научных конференций, симпозиумов, семинаров, школ[5].

Всем текстовым объектам, помимо атрибутов, что наследуются из вышестоящих объектов, присущи свойства:

- язык контента;
- количество страниц (или диапазон) в опубликованной версии.

Эти объекты могут также иметь такие атрибуты как:

- содержание;
- набор файлов, когда объекты данного типа располагают полным текстом, и он хранится в файлах,

а также связи с аналогичными объектами:

- является переводом;
- имеет перевод;
- является версией;
- имеет версию;
- цитируется;
- цитирует.

Этот перечень связей может быть существенно дополнен связями, рассмотренными в модели FRBR.

Каждый объект, имеющий тип *словесность*, обладает своими присущими только ему атрибутами или связями. Например, *книга*, помимо перечисленных атрибутов для вышестоящих в иерархии физического, рукотворного объекта и объекта словесность также имеет:

- ISBN (Международный стандартный номер книги);
- издатель;
- издание;
- место и дата публикации;
- автор предисловия (послесловия);
- редактор;
- автор перевода (если она переводная).

Изданная статья помимо общих атрибутов имеет: обязательный атрибут-связь *выпуск*, связывающий экземпляры класса *изданная статья* с соответствующим экземпляром класса *выпуск*. Объект *выпуск* (имеется ввиду журнала), который мы также отнесли к классу *словесность,* дополнительно имеет атрибуты *дата*, *номер*, может иметь *том* и *тему выпуска*, а также связующие свойства *журнал* и *изданная статья*.

Обсуждая иерархию объектов, нужно также перечислить классификаторы, используемые при задании некоторых их атрибутов. В модели CIDOC CRM эта категория сущностей выделена в *концептуальный объект* (рис. 1). Так, например, атрибуты *тема* и *ключевые слова* как правило задаются с помощью распространенных тематических или предметных классификаторов: УДК, ББК, тематического классификатора ВАК и некоторых других: DDC, LCC, LCSH, MESH. Атрибут *язык* желательно определять в соответствии со стандартами RFC 1766 (ISO 639-2, ISO 3166); *географическое положение* – в стандарте GEO; форматы файлов задавать контролируемым словарем MIME; при описании сущности *человек* использовать набирающий все большую популярность FOAF.

3.2 Связи

Выше в описании иерархии объектов и их свойств уже упоминались некоторые связи, важной составляющей концептуальной модели и любой ЭБ. Объекты информационного пространства ЭБ связаны между собой бинарными ориентированными обязательными связями вида "1..1" (один к одному) или "1..*" (один ко многим) и необязательными связями вида "0..1" или "0..*". Примеры таких связей показаны на рис. 6.

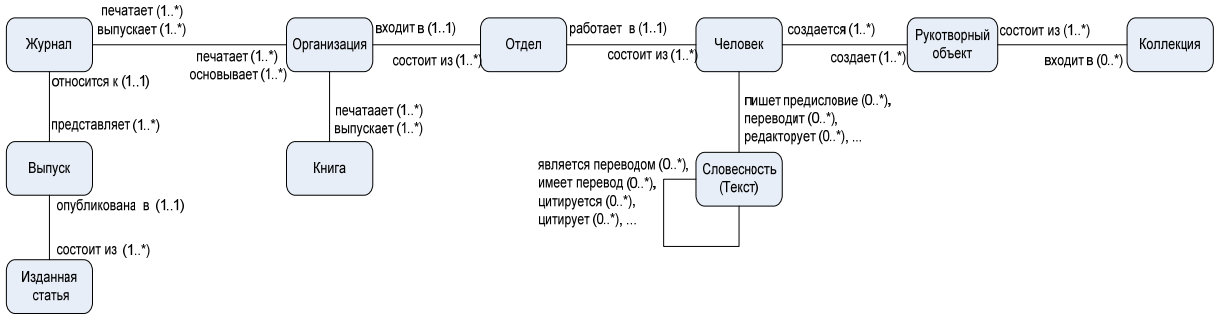

Рис. 6. Связи между сущностями в информационной модели научной электронной библиотеки

Заключение

В мире электронных библиотек существуют задачи, которые могут быть решены с помощью качественного концептуального описания таких систем. На сегодняшний день не существует всеохватывающей модели ЭБ, которую можно было бы с полным правом назвать эталонной. Построение хорошей модели, учитывающей мировой опыт подобных разработок, является нашей целью.

В данной работе дан краткий обзор трех известных проектов из этой области: CIDOC CRM, FRBR, DELOS DLRM, а также рассмотрена информационная составляющая разрабатываемой нами концептуальной модели ЭБ по состоянию на текущий момент. Пользовательская и функциональная составляющие модели ЭБ находятся в зачаточной стадии разработки и пока что не представлены. Авторы надеются на интересное и полезное обсуждение.

Литература

1. *Когаловский М.Р., Паринов С.И.* Информационные ресурсы, наукометрические показатели и показатели качества метаданных системы Соционет // Труды Девятой Всероссийской конференции "Электронные библиотеки: перспективные методы и технологии, электронные коллекции" — RCDL'2007, г. Переславль-Залесский, Россия, 15-18 октября 2007 г. — С.45-54.
2. *A Guide* to Institutional Repository Software. 3rd Edition. Open Society Institute. 2004. http://www.soros.org/openaccess/pdf/OSI_Guide_to_IR_Software_v3.pdf
3. *Candela L., Castelli D., Fuhr N., Ioannidis Y., Klas C.-P., Pagano P., Ross S., Saidis C., Schek H.-J., Schuldt H., Springmann M.* 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.
4. *Crofts N., Doerr M., Gill T., Stead S., Stiff M. (editors),* Definition of the CIDOC Conceptual Reference Model, January 2008. Version 4.2.4
5. *Functional Requirements for Bibliographic Records, Final Report* / IFLA Study Group on the Functional Requirements for Bibliographic Records. – München: K.G. Saur, 1998. (UBCIM Publications, New Series; v. 19) http://archive.ifla.org/VII/s13/frbr/frbr.htm
6. *Candela L., Castelli D., Dobreva M., Ferro N., Ioannidis Y., Katifori H., Koutrika G., Meghini C., Pagano P., Ross S., Agosti M., Schuldt H., Soergel D.* The DELOS Digital Library Reference Model Foundations for Digital Libraries. IST-2002-2.3.1.12. Technology-enhanced Learning and Access to Cultural Heritage. Version 0.98, December 2007. http://www.delos.info/files/pdf/ReferenceModel/DELOS_DLReferenceModel_0.98.pdf
7. *The CIDOC Conceptual Reference Model. http://cidoc.ics.forth.gr*
8. *Doerr M., Iorizzo D.* The Dream of a Global Knowledge Network — A New Approach // ACM Journal on Computing and Cultural Heritage, Vol. 1, No. 1, Article 5, Publication date: June 2008.
9. *Tillett*, *Barbara B.* "Bibliographic Relationships." In: *Relationships in the Organization of Knowledge*, edited by Carol A. Bean and Rebecca Green. — Dordrecht : Kluwer Academic Publishers, 2001, p. 19-35.
10. *Barbara Tillett* What is FRBR? A Conceptual Model for the Bibliographic Universe. http://www.loc.gov/cds/downloads/FRBR.PDF
11. *Функциональные* требования к библиографическим записям : окончат. отчет / Рос. библ. ассоц., Рос. гос. б-ка ; пер. с англ. [В.В. Арефьев ; науч. ред. пер.: Т.А. Бахтурина, Н.Н. Каспарова, Н.Ю. Кулыгина]. — Москва : РГБ, 2006. — [150] с.
12. *Doerr M., Leboeuf P.* Modelling intellectual processes: The FRBR—CRM harmonization // In Conference Proceedings of ICOM-CIDOC Annual Meeting. Gothenburg, Sweden. — 2006. — pp. 10–14.

Conceptual Model of Digital Library

Reznichenko V.A., Proskudina G.Yu, Kudim K. A.

The paper is concerned with creation of scientific digital library conceptual model. Several well-known connected projects are surveyed, those are CIDOC CRM, FRBR, DELOS DLRM. Informational part of a conceptual model being developed is presented.

[1] http://dspace.nbuv.gov.ua:8080/dspace
[2] http://eprints.isofts.kiev.ua/
[3] http://www.greenstone.org/
[4] http://www.fedora-commons.org/
[5] http://www.opendlib.com
[6] Подробный перечень типов документов представлен, например, в одной из вспомогательных таблиц «Формы документов» классификатора УДК.