

**FRAD – одна из трёх моделей представления метаданных
о документах в информационных системах**

Описание модели «Функциональные требования к авторитетным данным» (FRAD), разработанной международной рабочей группой FRANAR (ИФЛА), издано на английском языке в 2009 г. Перевод на русский язык, выполненный автором доклада, опубликован РБА в 2011 г. в виде книги и размещён на своём сайте и сайте ИФЛА.

Модель FRAD рассматривается в контексте двух других моделей той же серии – FRBR и FRSAD. Представлены подробное описание концептуальной основы и структуры FRAD, а также некоторые терминологические проблемы перевода.

Ключевые слова: модель «Функциональные требования к авторитетным данным», нормативные данные, библиографические объекты, идентификаторы, контролируемые точки доступа.

Взаимосвязанные модели данных

Три модели представления данных о документах в электронной среде появились на свет в результате творческой деятельности трёх рабочих групп, созданных ИФЛА, и обсуждения на международном уровне. Важным обстоятельством, с нашей точки зрения, является то, что эти модели могут быть использованы не только в электронных каталогах, но и для создания иных формальных способов формирования метаданных в электронных библиотеках и архивах, музейных информационных системах, других мультимедийных базах данных.

Речь идет о следующих моделях:

- 1) *Functional Requirements for Bibliographic Records, FRBR* [1] (Функциональные требования к библиографическим записям [2]) – рабочая группа FRBR;
- 2) *Functional Requirements for Authority Data, FRAD* [3] (Функциональные требования к авторитетным данным [4]) – рабочая группа FRANAR;
- 3) *Functional Requirements for Subject Authority Data, FRSAD* [5] (Функциональные требования к предметным авторитетным данным) – рабочая группа FRSAR.

В самих названиях моделей отражён процесс развития взглядов на уровень их абстракции: сначала моделировали структуру библиографических записей (FRBR), затем создали рабочую группу под названием «Функциональные требования к авторитетным/нормативным записям и их нумерация» (FRANAR), которая и сделала вывод, что более перспективны модели данных вне зависимости от конкретного способа их представления (не обязательно будут использоваться записи в формате типа MARC).

Действительно, в последние годы в свои права вступило, в частности, представление метаданных с помощью специальных языков разметки (например XML и его варианты), развиваются способы трёхмерного изображения связей между структурными элементами моделей. Из этих соображений группа FRANAR, в которой довелось работать и автору доклада, заявила о подготовке модели не для записей, а для данных: в рамках этого исследования «не делались а priori некие допущения ни относительно физической структуры авторитетных данных, ни о способах хранения данных в форме авторитетного файла, существующего отдельно от каталога или библиографического файла как такового или полностью интегрированного с ним» [4. С. 14]. Затем и группа FRSAR, в названии которой ещё присутствовало слово «записи», поступила аналогичным образом.

Дискуссия по определению уровня абстракции модели FRAD длилась несколько лет. В ходе международного обсуждения промежуточных вариантов документа одни специалисты упрекали

разработчиков в том, что модель далека от библиотечной практики, другие, напротив, в том, что модель должна быть приближена к уровню семиотических и лингвистических исследований.

В результате рабочая группа остановилась на некотором компромиссном уровне; но особенно важно, что удалось сделать *ключевым понятием модели «имя»* в лингвистическом и семиотическом значении термина, добавив в модель термин «идентификатор».

Рабочая группа – FR SAR ввела ещё более абстрактное понятие с использованием термина «*номен*», в объём содержания которого входят «*имя*» и «*идентификатор*». Различие между именем и идентификатором одно: имя может ассоциироваться не только с одним объектом, а идентификатор – только с одним.

Особое значение модели *FRBR* состояло в том, что она обратила внимание специалистов на следующий факт: библиографические объекты не являются сущностями некоторого одного порядка и, следовательно, не могут быть описаны одними и теми же методами, по единой модели. Это объясняется тем, что они могут рассматриваться как сущности (или объекты) различных уровней.

К сожалению, не удалось ввести в систему терминологии модели *FRAD* термин «*метаданные*». Представляется, что это понятие хорошо структурирует взаимосвязи между тремя моделями.

Известно, что библиографические данные формируются как метаданные по отношению к данным на титульном листе документа или к данным смыслового характера, отраженным средствами естественного языка в тексте документа. Это область модели *FRBR*, которая выделяет и описывает следующие основные сущности (объекты): произведение, выражение, воплощение и физическая единица, а также: лицо, организация, концепция, предмет, событие, место. В контексте описываемых моделей их называют библиографическими объектами. Имена рассматриваются в *FRBR* как атрибуты этих объектов (имя лица, наименование организации, заглавие воплощения). Данные, описывающие смысловое содержание документа (предмет, тема), не рассматриваются.

Модели *FRAD* и *FRSAD* структурируют метаданные следующего уровня, которые предназначены для контроля точек доступа в библиографических данных для задач информационного поиска [6, 7]. Они рассматривают в качестве своих объектов только данные, являющиеся контролируемыми точками доступа, обеспечивая их нормативный (авторитетный) контроль. Этот контроль означает как идентификацию объектов, представленных в качестве контролируемых точек доступа, так и непрерывный процесс управления ими [4].

Указанные модели основаны на модели *FRBR* и разрабатывались практически параллельно. Все три модели предложено [5. С. 45] обозначать как «*семейство FRBR*» (*FRBR family*), поскольку они представляют собой как бы части одной более широкой модели.

С нашей точки зрения, пора переименовать модель *FRBR* в модель данных (не записей) – *FRBD*. Предполагая, что такого рода общая модель будет создана, разработчики *FRSAD* провели сравнительный анализ трёх моделей [5, приложение В] и выявили различия между ними, требующие гармонизации.

Модель *FRAD* структурирует данные об именах, добавляя наряду с ними понятие «*идентификаторы*» и рассматривая те и другие в рамках данной модели в качестве сущностей (объектов), а не атрибутов. При этом авторитетные/нормативные данные «определены как совокупность информации о лице, роде/семье, организации или произведении, имена которых используются как основа для контролируемых точек доступа» [4] к библиографическим данным. Однако *FRAD* не касается сущностей, которые связаны с произведением отношениями «являться предметом» или «иметь предметом», т.е. предметных (тематических) связей между произведением и любыми другими сущностями (объектами), которые являются его предметами.

Модель *FRSAD* структурирует преимущественно данные, которые имеют прямое отношение к семантической стороне произведения. Здесь появляются термины «*thema*» (слово из латинского языка, отсутствующее в английском) и «*номен*» (также из латыни) [5], которые, не мудрствуя лукаво, имеет смысл перевести как «*тема*» и «*номен*».

На самом деле проблема перевода терминов здесь не столь тривиальна, как кажется, поскольку русские термины требуется также «выстроить в ряд» без потери смысла и без повторения слов.

Получилось следующее:

1) *object - entity – subject – thema – work*

предмет (как материальная вещь) – объект (удачнее было бы – сущность) – предмет произведения – тема – произведение;

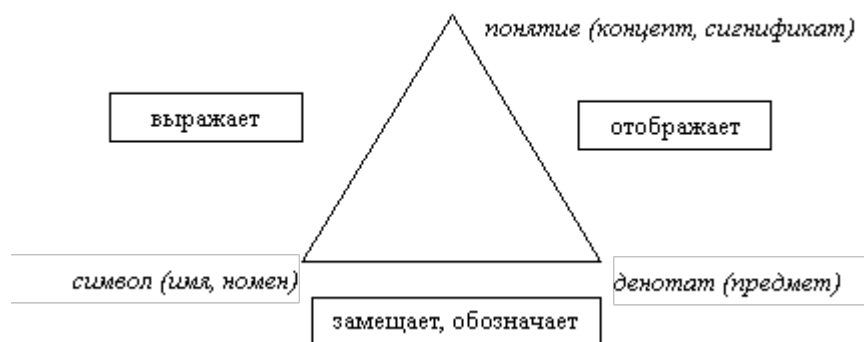
2) *name – identifier – nomen*

имя – идентификатор – номен.

«**Тема**» представлена в модели FRSAD [5] в качестве суперкласса для всех других сущностей (*объектов, entities*), которые могут стать предметами произведений (*subjects of works*). В этом смысле можно считать, что данные, описываемые в этой модели, представляют собой более высокий уровень метаданных. Это требует дальнейшего осмысления.

«**Номен**» определен как суперкласс (т.е. более широкое понятие) для сущностей «*имя*», «*идентификатор*» и «*контролируемая точка доступа*» из модели FRAD, которые рассматриваются как типы номена. Заметим, что упоминание «контролируемой точки доступа» в этом ряду представляется излишним. Номен рассматривается в модели и как сущность, и как атрибут.

Отношение «тема–номен» составляет основу модели FRSAD, причём делается ссылка на знаменитый семиотический треугольник:



Имя как таковое предназначено для именования предмета окружающего мира и поэтому находится в неразрывной связи с денотатом (предметом) и соответствующим сигнификатом (понятием, концептом, десигнатом), т.е. имя обозначает предмет и выражает понятие.

В модели такое же место занимает **номен** (имя, идентификатор), определяемый как знак или последовательность знаков, а прочие сущности (объекты) и сами имена и идентификаторы могут стать семантической составляющей для номена. С лингвистической точки зрения получается, что отношение «быть предметом» соединяет смысловую сторону «предмета», т.е. тему, с его материальной стороной – номеном. Это вполне укладывается в понимание предмета произведения как знака (лексической единицы) информационно-поискового языка (классификационного, дескрипторного, предметных рубрик).

Контроль имен, в широком смысле этого термина, и идентификаторов с середины прошлого века происходит в автоматизированных информационных системах на базе специальных словарей, которые выполняют нормативную функцию, определяя нормы использования имен (терминов, предметных рубрик, дескрипторов, индексов классификации) и их не принятых в системе вариантов, которые могут быть введены пользователем при поиске, а также словарей (номенклатур) идентификаторов. Отсюда распространенные словосочетания *нормализация лексики, нормативные словари, нормативный контроль, нормативная информация, нормативные данные* и т.д.

Эти словари (файлы) задают формы имен и идентификаторов, их атрибуты, необходимые и достаточные для отождествления имени с обозначаемым им предметом и понятием, связи между именами, предметами и понятиями. Естественно, нормативные данные в аспекте информационного поиска реализуются в виде

нормативных словарей, имеющих различные физические формы. Для представления такого рода данных и служат модели FRAD и FRSAD.

Структура модели FRAD

Итак, нормативные данные работают для обеспечения контроля имен и идентификаторов, используемых как контролируемые точки доступа в библиографических записях для обеспечения основных задач пользователей, которые определены в модели FRAD: найти, идентифицировать, контекстуализировать объекты и обосновать их выбор. Роль различных элементов модели для реализации этих задач подробно освещена [3, 4. Раздел 6].

Фундаментальная основа концептуальной модели изображена на рисунке [4.С. 17]. Модель можно описать в простой форме следующим образом: объекты в библиографическом универсуме (например идентифицированные в FRBR) известны под именами и/или идентификаторами. В процессе каталогизации (это может происходить в библиотеках, музеях или в архивах) эти имена и идентификаторы используются как основа для построения контролируемых точек доступа.

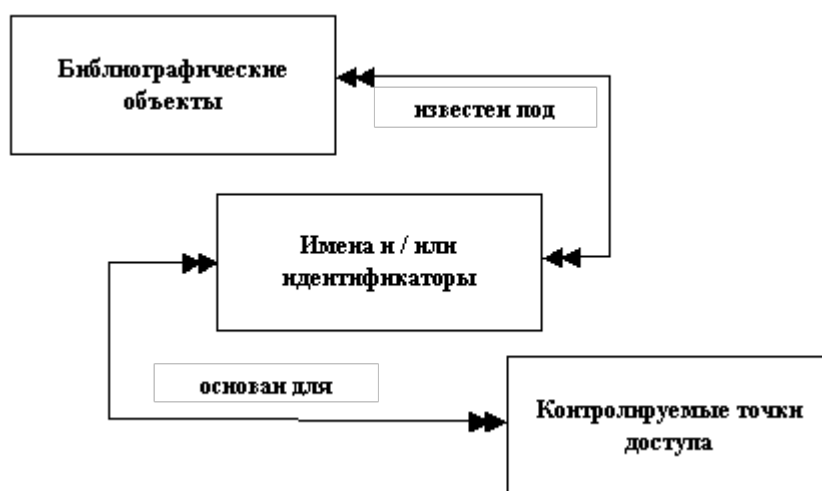


Рис. Фундаментальная основа концептуальной модели FRAD

Основу каждого вида данных составляет выбор сущностей (или объектов рассмотрения, изучения), затем выявление их атрибутов и связей между сущностями (объектами).

На наиболее высоком уровне концептуальная модель FRAD включает нормативные данные всех типов. Диаграмма «объект–связь» и определения объектов [4. Раздел 3.4] отражают нормативные данные для лиц, родов/семей, организаций и географических объектов или для объектов, представленных в каталогах как нормативные данные для заглавий (различные типы заглавий произведений или произведений/выражений и обобщающих заглавий произведений или произведений/выражений), а также данные для «авторов – заглавий», для предметных нормативных данных (предметные термины, термины тезауруса и индексы классификации), нормативные данные для различных типов имён и идентификаторов (стандартных номеров, кодов и т.д.), связанных с этими объектами. Однако детальное исследование атрибутов объектов и связей, ассоциируемых исключительно с предметными нормативными данными, в описании модели FRAD не содержится.

Концептуальная модель нормативных данных [4] отражает связи между: библиографическими объектами (лицо, род/семья, организация, произведение, выражение, воплощение, физическая единица, концепция, предмет, событие и место) и их именами и идентификаторами; лицом, семьей/родом или организацией и произведением, выражением, воплощением или физической единицей; именами и идентификаторами и контролируруемыми точками доступа; контролируемыми точками доступа и правилами; контролируемыми точками доступа и учреждениями; правилами и учреждениями.

Конкретный пример одного из библиографических объектов может быть известен под одним или более именами (связь «иметь имя») и, наоборот, одно имя может ассоциироваться с одним или несколькими конкретными примерами какого-либо библиографического объекта (связь «являться именем для»).

Аналогичным образом конкретный пример одного из библиографических объектов может быть связан с одним или несколькими идентификаторами (*связь «обозначен»*), но каждый идентификатор может быть связан только с одним или неконкретным примером библиографического объекта (*связь «присвоен»*) [4].

Между лицом, семьей/родом или организацией, с одной стороны, и произведением, выражением, воплощением или физической единицей, с другой, может существовать *связь «ассоциируется с»*.

Конкретное имя или идентификатор может быть основой для контролируемой точки доступа (*связь «основа для»*) и, наоборот, контролируемая точка доступа может быть основана на имени или идентификаторе (*связь «основана на»*). Контролируемая точка доступа может быть основана на комбинации двух имен и/или идентификаторов, как это делается для точки доступа «создатель–заглавие», представляющей произведение и сочетающей в себе имя автора и имя (т.е. заглавие) произведения.

Контролируемая точка доступа может определяться с помощью правил (*связь «определяется с помощью/определяет»*), и эти правила могут, в свою очередь, использоваться одним или более учреждениями (*связь «применяются/применяет»*). Аналогичным образом контролируемая точка доступа может быть создана или модифицирована одним или более учреждениями (*связи «создана/создает» и «модифицирована/модифицирует»*).

В работе приведены определения объектов, описаны атрибуты объектов (лица, организации, произведения, имени, контролируемой точки доступа и т.д.).

Для наглядности приведем пример определения объекта «идентификатор» [4. С. 35]: «Число, код, слово, фраза, логотип, эмблема и т.д., которые ассоциируются с объектом и служат для дифференциации этого объекта от других объектов в том поле деятельности (сфере знаний), в котором присвоен идентификатор. Идентификатор может состоять из цепочки знаков идентификатора (т.е. из последовательности чисел и/или букв, присвоенных объекту для того чтобы служить уникальным идентификатором в том поле деятельности, в котором он присвоен) и префикса или суффикса (знака или набора знаков (чисел и/или букв), стоящих перед или после цепочки знаков идентификатора)».

Объект «идентификатор» включает такие идентификаторы, как номера полисов социального страхования, присвоенных государственными органами; персональные идентификаторы, присвоенные другими регистрационными службами; коммерческий регистрационный номер, регистрационный номер благотворительной организации и т.д., присвоенный государственным органом; идентификатор организации, присвоенный регистрационной службой (например ISBN, присваиваемый издательством); стандартный идентификатор, присвоенный регистрационной службой для идентификации содержания (например ISRC, ISWC, ISAN); стандартный идентификатор, присвоенный регистрационной службой для идентификации воплощений, т.е. изданий (например ISBN, ISSN); порядковый номер по тематическому каталогу или каталогу издателя, который присваивается музыкальному произведению музыковедом или издателем; индексы систематического каталога; идентификаторы для физических единиц, присвоенные хранилищами (например полочный индекс); классификационный индекс, установленный для конкретного объекта (например классификационный индекс, созданный для отдельного произведения живописи); зарегистрированная торговая марка.

Некоторые из этих идентификаторов могут быть включены в авторитетные данные, но при условии соблюдения конфиденциальности их хранения как личных данных.

Эти идентификаторы могут относиться только к объектам и не могут быть номерами авторитетных записей.

Приведём пример описания атрибутов имени как объекта [4. С. 54]:

Тип имени	Категория имени. Включает имена лиц, наименования организаций, родовые имена, торговые марки и заглавия произведений и воплощений. Включает имена концепций, предметов,
------------------	---

	событий и мест.
Последовательность знаков имени	Цепочка цифр и/или букв или символов, которая изображает имя объекта. Форма произведения, ассоциируемая с конкретным именем персоны.
Область применения	Включает формы, жанры и т.д. (например, литературные произведения, критические работы, труды по математике, детективные романы), ассоциируемые с именем, используемым автором.
Даты использования	Даты, ассоциируемые с использованием конкретного имени, установленного для лица, организации или рода/семьи.
Язык имени	Язык, на котором имя выражено.
Графика имени	Графика, в которой представлено имя.
Система транслитерации имени	Система, используемая для формирования транслитерированной формы имени.

В работе подробно анализируются связи между различного рода объектами модели, примерами одного рода объектов, именами объектов одного рода и т.д.

Пример таблицы связей между различными именами лиц, родов/семей, организаций и произведений [4. С. 84]:

Тип имени	Образцы типов связей
Имя лица <-> Имя лица	<ul style="list-style-type: none"> • связь «прежнее имя» • связь «более позднее имя» • связь «альтернативная лингвистическая форма» • связи «другой вариант имени»
Имя рода/семьи <-> Имя рода/семьи	<ul style="list-style-type: none"> • связь «альтернативная лингвистическая форма»
Имя организации <-> Имя организации	<ul style="list-style-type: none"> • связь «полное имя» • связь «акроним/инициалы/сокращения» • связь «альтернативная лингвистическая форма» • связи «другой вариант имени»
Имя произведения <-> Имя произведения	<ul style="list-style-type: none"> • связь «альтернативная лингвистическая форма» • связь «условное имя» • связи «другой вариант имени»

Конкретные примеры связей [4]:

Связь «прежнее имя» (имена лиц). Связь между именем лица и именем, которое лицо использовало в более ранние периоды жизни.

Связь между именем лица «Charlotte Nichols», полученным в результате брака с господином Николсом (Nichols), и прежним именем «Charlotte Bronte».

Связь «альтернативная лингвистическая форма» (имена организаций). Связь между именем организации и альтернативной лингвистической формой имени, под которой она известна. Сюда относятся и переводы имени организации.

Связь между именем организации, известной как «United Nations Organization» - на английском языке, «Organisation des Nations Unies» – на французском языке и «Организация Объединенных Наций» – на русском языке.

В конце оригинала книги [3] приведен указатель терминологии. В отличие от оригинала, термины в указателе к русскому переводу располагаются в порядке русского, а не английского алфавита. При этом для каждого термина на русском языке указывается его аналог на английском языке из оригинала.

Перспективы

Модель FRBR во многом определила тенденции современного развития способов представления библиографических данных в информационных системах. Можно предположить, что модель FRAD окажет серьезное влияние на взгляды относительно авторитетных/нормативных данных для контроля точек доступа к документам в традиционных и электронных библиотеках, в том числе на интерпретацию и перевод используемой в различных языках терминологии. В частности, уже внесены изменения в ряд публикаций ИФЛА.

Надо полагать, перевод книги на русский язык [4] будет способствовать успешному обсуждению и использованию модели FRAD российскими специалистами.

Список источников

1. **Functional** requirements for bibliographic records : final report / IFLA Study group on the Functional Requirements for Bibliographic Records. – München : K. G. Saur, 1998.
2. **Функциональные** требования к библиографическим записям : окончательный отчет / Рос. библ. ассоц., Рос. гос. б-ка, Рабочая группа ИФЛА по Функциональным требованиям к библиогр. записям ; [науч. ред.: Т. А. Бахтурина, Н. Н. Каспарова, Н. Ю. Кулыгина] ; пер. с англ. [В. В. Арефьев]. – Москва : Пашковдом, 2010 – 165 с.
3. **Functional** Requirements for Authority Data – A Conceptual Model. // Series: [IFLA Series on Bibliographic Control](#) 34. – München: K. G. Saur, 2009. – 101 p.
4. **Функциональные** требования к авторитетным данным : концептуальная модель : заключительный отчет, декабрь 2008 / под ред. Гленна Е. Патона ; Рабочая группа ИФЛА по разработке функциональных требований к авторитетным записям и их нумерации (FRANAR) ; [пер. с англ.: О. А. Лаврёнова]. – С.-Петербург : Изд-во «Российская национальная библиотека», 2010. – 115 с. – Режим доступа: <http://www.ifla.org/en/publications/functional-requirements-for-authority-data>
5. **Functional** Requirements for Subject Authority Data. –Режим доступа:
<http://www.ifla.org/node/1297>
6. **Лавренова О. А.** Точки доступа при поиске и их контроль // Сайт ЭЛБИ. Блоги. – Режим доступа: <http://www.aselibrary.ru/blogs/?p=375>
7. **Лавренова О. А.** Точки доступа к электронным ресурсам и модель нормативных данных FRAD // Библиотечные, музейные, архивные учреждения в век электронных коллекций и библиотек : мат. VI Науч.-практ. семинара «Электронные ресурсы библиотек, музеев, архивов», 28–29 окт. 2010 г.,

Санкт-Петербург / ЦГПБ им. В. В. Маяковского ; ред.-сост. И. Е. Прохоров. – С.-Петербург : Северная звезда, 2010. – С. 187–196.