

Этот доклад читался на конференции [«Современные информационные технологии и письменное наследие: от древних текстов к электронным библиотекам»](#) [[Modern Information Technologies and Written Heritage: From Ancient Texts to Electronic Libraries](#)] (E1Manuscript-08), которая произошла 25–30 августа 2008 в городе Казань, Татарстан, Российская Федерация. Английский оригинал тоже доступен в интернете.

Логическое следование сущностей и импликатура атрибутов в ER-модели «Функциональных требований к библиографическим записям»¹

Кевин С. Хокинс (Kevin S. Hawkins)

Мичиганский университет, г. Энн-Арбор, Мичиган, США (University of Michigan, Ann Arbor, Michigan, USA)

Введение

Исследователи, занимающиеся кодированием текстов, много спорят о том, может ли быть сформулирована модель данных для текстов. Использование компьютеров в библиотеках началось тогда же, когда и их применение для кодирования текстов, но лишь недавно библиотекари стали пытаться сформулировать собственную модель данных для наиболее распространенного вида применения компьютеров в библиотеках: для библиотечных каталогов.

Однако наиболее часто упоминаемая в библиотечных исследованиях модель – модель, описанная в работе «Функциональные требования для библиографических записей: Окончательный отчет» (Functional Requirements for Bibliographic Records: Final Report, далее «FRBR») – это не полная модель данных, что признается и в самом отчете, и в его критике. Тем не менее, она часто воспринимается как модель данных, причем обсуждается иерархия сущностей и наследование свойств в этой модели, несмотря на то, что ни одно из этих понятий не упоминается в самом отчете.

В настоящей статье указывается на неадекватность FRBR как модели данных и приводятся аргументы в пользу того, что в данной модели должен быть расширен функциональный подход для обеспечения возможности машинного обучения библиографической информации, отчасти на основании строгого наследования, но большей частью на основании нормативности и имплицативных допущений.

«Функциональные требования для библиографических записей: Окончательный отчет»

Отчет FRBR был впервые опубликован Международной федерацией библиотечных ассоциаций и учреждений (International Federation of Library Associations and Institutions (IFLA)) в 1998 году. Одна из двух его целей – «создать ясно очерченную, структурированную систему для того, чтобы поставить данные, хранящиеся в библиотечных записях, в соответствие с нуждами пользователей этих записей» [IFLA 2008: 7]. Изначально целью было повышение способности каталогов к взаимодействию, однако с течением времени эта идея стала восприниматься как попытка полностью реорганизовать библиотечные каталоги, уничтожая последние следы библиотечных карточек.

Обсуждение FRBR в основном касается сущностей «группы 1» – произведение, форма выражения, проявление и физическая единица. (Ниже эти термины будут использоваться только в том же смысле, что и в FRBR). Определения и примеры даны в таблице.

¹ Автор благодарен Ольге Хомицевич за перевод на русский язык (с помощью перевода Т.А. Бахтурины терминов FRBR) и очень признателен Дэвиду Дубину, Эллену Риниру, и Ричарду Урбану за помощи в развитии доклада.

Сущность	Определение	Примеры
произведение (work)	отдельное интеллектуальное или творческое произведение (“a distinct intellectual or artistic creation” [IFLA 2008: 17])	«Мастер и Маргарита» М.В.Булгакова
форма выражения (expression)	интеллектуальная или творческая реализация произведения в виде текстовой, музыкальной или хореографической записи, звука, изображения, объекта, движения и т.п., или любое сочетание этих видов записи (“the intellectual or artistic realization of a work in the form of alpha-numeric, musical, or choreographic notation, sound, image, object, movement, etc., or any combination of such forms” [IFLA 2008: 19])	текст первой версии романа, который Булгаков сжег; цензурированная версия, опубликованная в журнале «Москва» в 1966-1967 гг; перевод на английский язык Майкла Гленни
проявление (manifestation)	физическое воплощение формы выражения произведения (“the physical embodiment of an expression of a work” [IFLA 2008: 21])	перевод Гленни, опубликованный в мягкой обложке издательством Harper & Row; перевод Гленни, опубликованный как аудиозапись
физическая единица (item)	одиночный экземпляр проявления (“a single exemplar of a manifestation” [IFLA 2008: 24])	мой экземпляр перевода Гленни, изданный в мягкой обложке в издательстве Harper & Row

Следует отметить, что FRBR намеренно рассматривает и не книжные носители информации.

Сущности, относимые FRBR к группе 1, часто понимаются как иерархия, на вершине которой находится произведение, а внизу – физическая единица (см., например, [Mimno и др. 2005]), однако FRBR не дает однозначного ответа на то, являются ли эти сущности иерархически организованными² и, более того, было отмечено, что идея иерархии проблематична для FRBR [Coyle 2007, Renear и Choi 2006]. Тем не менее, в нетерминологическом смысле это удобное обозначение для этих сущностей.

Кроме этих сущностей, а также других, которые здесь не рассматриваются, модель FRBR также определяет свойства. Только определенные свойства относятся

² Термин «иерархический» употребляется в FRBR только один раз, во введении: «Могут быть проведены дальнейшие исследования в области практического применения реструктуризации MARC-форматов для более точного отражения иерархических и взаимных отношений, описанных в данной модели.

к определенным сущностям; некоторые свойства, такие как название, существуют для более чем одной сущности, однако это, строго говоря, разные свойства, которые случайно оказываются омонимичны. Например:

1. Произведение, форма выражения и проявление могут иметь названия, но эти названия не обязательно совпадают.
2. Проявление может иметь определенный шрифт, но это свойство не относится к другим сущностям.
3. Форма выражения может иметь ограничения на использование, в то время как проявление и физическая единица могут иметь ограничения на доступ.

Свойства, описанные в FRBR, еще более очевидным образом происходят из терминологии, относящейся к области каталогов, чем описанные там сущности, и они в значительной степени отражают практику составления каталогов и нужды библиотек. К примеру, название произведения – это унифицированное заглавие в каталогах, в то время как название проявления – это текст, напечатанный на титульном листе. Другой пример – ограничения на использование и ограничения на доступ: ограничения на использования обычно связаны с копирайтом, а ограничения на доступ применимы к проявлению (в случае технологических мер по борьбе с пиратством) или к физической единице (в случае, если данная физическая единица может быть использована только в пределах библиотеки). Такой сугубо функциональный подход – причина популярности модели в библиотечном сообществе, однако он также приводит к ее неадекватности в качестве формальной модели данных.

FRBR как модель данных

Обладая обширной системой сущностей, свойств и отношений, FRBR является основой для модели данных библиографической вселенной. Однако эта модель неполна. Спорный вопрос об иерархии – не единственный случай, когда модель данных, лежащая в основе FRBR, недостаточно определена. FRBR признает, что использует анализ сущность-связь [IFLA 2008: 6], но прямо утверждает, что концептуальная модель отчета «не доводит анализ до того уровня, который требуется для полноценной модели данных» [IFLA 2008: 3]. Ученые Ринир и Чой перечисляют трудности, которые связаны с приведением модели FRBR к виду строгой модели сущность-связь. [Renear и Choi 2006].

Исследователи, заинтересованные в формулировании более строгой версии модели FRBR, которая могла бы быть использована компьютерными приложениями в качестве формальной онтологии, разрабатывают объектно-ориентированное определение FRBR под названием FRBR_{OO} [International 2008], которое представляет собой попытку примирить FRBR с существующей в музейном сообществе онтологией под названием «концептуальная справочная модель CIDOC» (CIDOC Conceptual Reference Model (CRM)) [ICOM/CIDOC 2007]. Модель CIDOC значительно более сложна, чем модель FRBR, однако она имеет ряд преимуществ, в том числе моделирование того, как появляются библиографические сущности.

Я согласен с тем, что FRBR_{OO} является гораздо более прочным основанием для развития настоящей модели данных библиографической вселенной. Однако я полагаю, что функциональный подход, свойственный исходной модели FRBR,

заслуживает внимания благодаря возможности улучшения стандартных способов использования библиотечных каталогов. В данной статье я предлагаю способы расширения модели FRBR для более точного соответствия ее восприятию и общепринятым представлениям о том, как библиографические сущности соотносятся друг с другом.

Как возникают сущности группы 1 по классификации FRBR?

Согласно FRBR₀₀, создается форма выражения, что приводит к созданию как произведения, так и «единицы проявления», что примерно соответствует физической единице в исходной модели FRBR [International 2008: 14]. Я согласен с идеей одновременной реализации сущностей, и поэтому предлагаю похожий подход для исходной модели FRBR. Продолжим рассмотрение примера романа «Мастер и Маргарита».

Булгаков писал этот роман на рукописных листах в 1928 и 1929 гг. Я предлагаю считать, что во время сочинения романа были представлены все четыре сущности группы 1 по классификации FRBR: произведение, форма выражения, проявление и физическая единица. Из акта творчества логически следует существование всех четырех сущностей.

Однако у многих произведений существует несколько форм выражения. В случае с «Мастером и Маргаритой», автор создал несколько версий, каждая из которых, согласно FRBR, является формой выражения. Более поздними формами выражения стали научные издания и переводы. При публикации формы выражения, она поступает в набор либо как-то иначе готовится к распространению; каждая отдельная версия для публикации – это новое проявление. Наконец, изготавливаются экземпляры этого проявления; каждый экземпляр – это физическая единица. Иерархия расширяется книзу, поскольку зачастую произведение имеет более чем одну форму выражения, форма выражения – более одного проявления, а проявлению соответствует более одной физической единицы. Все эти связанные между собой сущности составляют то, что принято называть «библиографическим семейством» [Taylor 2007: 73–78].

Следует заметить, что также могут возникать вторичные произведения – продолжения, работы поклонников и другие произведения, вдохновленные оригиналом. В данной статье они не рассматриваются.

На каком уровне иерархии применимы свойства FRBR?

Модель FRBR₀₀, как и исходная модель FRBR, имеет существенный недостаток: в ней не дается объяснения тому, как свойство присваивается сущности. Например, каким образом произведение, форма выражения или проявление получает название? Как соотносятся ограничения доступа к физической единице с ограничениями доступа к проявлению? Модель данных для библиографических сущностей нуждается в правилах логического следования, которые могут позволить компьютеру определить – или по крайней мере угадать – то, что не указано напрямую.

Как мне кажется, для пользователей будет удобнее, если мы будем рассматривать свойства FRBR как потенциально существующие на любом уровне иерархии, «унаследованные» от ранее существовавших сущностей внутри библиографического семейства, однако с возможностью изменения при

возникновении новой сущности. В таком случае свойства будут использоваться примерно так же, как и в повседневной жизни. Обратимся снова к Булгакову.

Булгаков написал черновик своего романа, создав при этом не только физическую единицу, но и проявление, форму выражения и произведение. Его творение имеет такие свойства, как название, автор и ограничения доступа, которые в одинаковой степени присутствуют на всех сущностях, хотя в этот момент в реальности существует только одна физическая единица.

Впоследствии, когда создаются дополнительные формы выражения, эти свойства передаются далее по иерархической лестнице, однако они могут измениться в любой ее точке. Возьмем, к примеру, название: издатель может изменить его (на уровне формы выражения), наборщик может поменять в нем написание с большой или маленькой буквы (на уровне проявления), владелец экземпляра книги может зачеркнуть его (на уровне физической единицы). В последнем случае имеет смысл говорить о заглавии физической единицы как о зачеркнутом, хотя FRBR не предполагает наличия свойства заглавия на уровне физической единицы.

Сходным образом дело обстоит с ограничениями на использование и доступ: ограничение на использование формы выражения (копирайт или лицензия) применимо ко всем проявлениям и физическим единицам, образованным от этой формы выражения, однако если проявление этой формы выражения позднее будет издано с лицензией Creative Commons, то измененное ограничение на использование будет применяться ко всем физическим единицам этого проявления.

Нормативность и импликация

Рассмотрение уровней, на которых существуют свойства, приводит к мысли о нормативности сущностей FRBR. Если в книге (физической единице) не хватает страницы, она все равно является физической единицей данного проявления, к которому тоже относятся другие экземпляры этого издания. Если издание (представление) публикуется под другим названием, чем предыдущие издания, оно все равно является реализацией формы данного выражения. В библиографической вселенной существует много информации, которую можно предположить, однако мало таких фактов, которые могли бы изменить наши основные предположения, если эти факты окажутся неверными.

Подобно работе Ринира и Чой, в литературе о FRBR часто говорится о «наследовании», подобном вышеописанной иерархической лестнице сущностей. Они утверждают, что наследование является недостаточным объяснением для FRBR, и предлагается использовать «модальные понятия» [Repear и Choi 2006]. Вместо этого я предлагаю использовать импликацию для описания поведения свойств. Импликация (implicature), термин, введенный лингвистом Полом Грайсом, означает умозаключение, которое может быть выведено, но не обязательно является верным. При существовании формы выражения, проявления или физической единицы, о которых нет полной информации, можно сделать вывод об их свойствах на основании уже имеющихся сущностей.

Очевидно, что импликация, в отличие от логического следования, не позволяет компьютеру выводить строгие заключения, а лишь позволяет предположить, что является возможным. Однако такие предположения могут быть очень полезны, и при этом они имитируют поведение человека при работе с

библиографическими семействами. При поиске библиографических сущностей, человек предполагает, что они имеют те же свойства, что и другие сущности в данном семействе, однако различия не вызывают удивления.

Заключение

Хотя очевидно, что FRBR не является самодостаточной моделью данных, она основывается не только на практике составления каталогов, но и на интуитивных представлениях о том, как сущности в библиографических семействах соотносятся друг с другом. Пользователи, ищущие библиографические сущности, строят предположения на основании известной информации, однако часто не могут сделать четких выводов. Если мы используем модель FRBR при создании инструментов для помощи пользователям в поиске информации, полезно будет применять логическое следование сущностей и импликацию свойств.

В частности, обновленная модель допускает наличие свойств практически на любом уровне модели FRBR. Я полагаю, что для функциональной модели, каковой является FRBR, может быть полезна еще менее строгая онтология, чем предлагается в самом отчете.

Библиография

- Coyle 2007 – Coyle, K. Hierarchy v. relationships [Электронный ресурс] / Coyle's InFormation : comments on the digital age, which, as we all know, is 42. – 2007. – Режим доступа: <http://kcoyle.blogspot.com/2007/11/use-of-hierarchy-as-organizing.html>, свободный. – Загл. с титул. страницы.
- ICOM/CIDOC 2007 – ICOM/CIDOC Documentation Standards Group. Definition of the CIDOC Conceptual Reference Model [Электронный ресурс] / produced by the ICOM/CIDOC Documentation Standards Group, continued by the CIDOC CRM Special Interest Group. – Version 4.2.2. – 2007. – Режим доступа: http://cidoc.ics.forth.gr/docs/cidoc_crm_version_4.2.2.pdf, свободный. – Загл. с титул. страницы.
- IFLA 2008 – IFLA Study Group on the Functional Requirements for Bibliographic Records. Functional requirements for bibliographic records : final report [Электронный ресурс] / IFLA Study Group on the Functional Requirements for Bibliographic Records; Approved by the Standing Committee of the IFLA Section on Cataloguing September 1997 – As amended and corr. through February 2008. – [2008]. – Режим доступа: http://www.ifla.org/VII/s13/frbr/frbr_2008.pdf, свободный. – Загл. с титул. страницы.
- International 2008 – International Working Group on FRBR and CIDOC CRM Harmonisation. FRBR object-oriented definition and mapping to FRBRER (version 0.9 draft) [Электронный ресурс] / International Working Group on FRBR and CIDOC CRM Harmonisation supported by Delos NoE ; editors Chryssoula Bekiari, Martin Doerr, Patrick Le Bœuf ; contributors Trond Aalberg ... [et al.] – 2008. Режим доступа: http://cidoc.ics.forth.gr/docs/frbr_oo/frbr_docs/FRBR_oo_V0.9.pdf, свободный. – Загл. с титул. страницы.
- Mimno и др. 2005 – Mimno, D., Crane, G., Jones, A. Hierarchical catalog records : implementing a FRBR catalog [Электронный ресурс] // D-Lib Magazine. – Vol.

- 11, no. 10. (Oct. 2005). Режим доступа:
<http://www.dlib.org/dlib/october05/crane/10crane.html>, свободный. – Загл. с
титул. страницы.
- Renear и Choi 2006 – Renear, A., Choi, Y. Modeling our understanding, understanding
our models : the case of inheritance in FRBR [Электронный ресурс] / Allen H.
Renear, Yunseon Choi. – [2006]. – Режим доступа:
http://eprints.rclis.org/archive/00008158/01/Renear_Modeling.pdf, свободный. –
Загл. с титул. страницы.
- Taylor 2007 – Taylor, A. Understanding FRBR : what it is and how it will affect our
retrieval tools / edited by Arlene G. Taylor. Westport, Conn. : Libraries Unlimited,
2007.