



**Сломать преграды между старыми методиками и новыми запросами: цена нерешительности**

**Maja Žumer  
Jan Pisanski  
Polona Vilar  
Viktor Harej  
Tanja Merčun  
Katarina Švab**

Department of Library and Information Studies and Book Science, Faculty of Arts  
University of Ljubljana  
Ljubljana, Slovenia

**Meeting:**

**80 — Cataloguing: breaking barriers — Cataloguing Section**

**Резюме:**

*За последнее время каталогизация претерпела немало основательных изменений. В настоящем докладе предложена оценка того, чего каталогизационное сообщество достигло к настоящему моменту, и предложение относительно того, на чем следует сконцентрироваться в будущем. Каталоги должны обеспечить читателям более удобный и целенаправленный способ достижения и использования информации, поскольку библиотеки обладают инструментами для раскрытия перед пользователями более широкого контекста, которого сегодня в каталогах нет. Как отмечено в двух наших исследованиях, у нас есть FRBR (а также - FRAD и FRSAD), которые мы могли бы и должны были бы использовать в качестве основы каталогизации. К настоящему времени на основе этих документов, а также – на основе Международных принципов каталогизации, разработаны некоторые своды правил каталогизации (возможно, наибольшее значение имеет RDA), представляющее собой важный шаг вперед.*

*Однако, можно определить несколько ключевых вопросов, на которых библиотечному сообществу необходимо сконцентрировать свое внимание, чтобы завершить процесс: гармонизация (различных процессов), FRBR -изация, идентификация и представление. Необходимо, чтобы каталогизация постоянно оставалась в центре каждой из этих областей, решения же должны приниматься на основе изучения пользователя, чем, к сожалению, до последнего времени занимались мало. Такая трансформация библиотечного сообщества слишком неспешна, и поэтому эти вопросы надлежит разрешить быстро и эффективно, - если библиотеки хотят сохранить свою роль и значение в мире сегодняшней информации.*

## **Революция на основе FRBR**

В последние годы каталогизация претерпевала немало революционных изменений, вызванных разработками и внедрением технологий, необходимостью изменить существующие строгие методики, которые как-то утратили свою задачу в оцифрованном пространстве, а также – последними изысканиями, связанными с информационными потребностями и стандартными поисковыми запросами пользователей. Несомненно, FRBR (Функциональные требования к библиографическим записям, 1998) – первая формальная модель библиографической среды – является для библиотечного сообщества и методик и результатов каталогизации дерзким шагом в некоем новом направлении

То, что FRBR принесёт выигрыш и пользователям, и библиотекарям, было предсказано сразу после обнародования исследования ИФЛА на одном из обучающих семинаров ИФЛА (Noerr, Goossens, Matei, Otten, Peruginelli и Witt, 1998): пользователи выигрывают за счёт более простого поиска, более точных результатов, группировки на уровне произведения, улучшенной навигации и большего удобства библиографических ссылок, тогда как библиотекарям выгодно то, что метаданные лучше размещаются в записях, каталогизация копированием и совместное использование записей упрощены, и можно добавлять новые метаданные (например, управление правами).

Когда дело доходит до предоставления информации, мы с грустью убеждаемся в том, что библиотечное сообщество не проявляло и не проявляет достаточной активности. Являясь держателем неоспоримо самого богатого источника информации о ресурсах, большинство библиотечных каталогов всё ещё основывается на модели, которая невероятно хорошо работала в карточных каталогах, но стала устаревшей с того самого времени, как впервые появились компьютеры. Как сформулировала Yee (2005), FRBR создает возможность развить он-лайнные списки общественных изысканий в подлинные он-лайнные каталоги общественного доступа.

### **Интуитивен ли FRBR?**

Хотя FRBR явно ориентировано на пользователя, основывалось оно не на целенаправленных изучениях пользователей (Madison, 2005). Поэтому многие призывали к тому, чтобы пригодность FRBR проверили пользователи, - вероятно, наиболее достойны упоминания Рабочая группа БК по будущему библиографического контроля

(2008) и специалисты по FRBR, упомянутые в исследовании Delphi, проведенном Zhang a Salaba (2009a). Этот пробел и вызвал наше исследование интуитивности FRBR. К настоящему времени нами выполнены два пользовательских исследования ментальных моделей библиографической среды. В обоих случаях мы рассматривали печатные книги.

Первое исследование, которое было проведено в 2007-2008 годах, занималось ментальными моделями лиц, не являющихся библиотекарями, использовавших три различных подхода: сортировка карточек, соответствие понятий и задача сравнения. Подробности исследования приводились неоднократно (Pisanski, Žumer, 2010a, 2010b) и его результаты показывают, что участники посчитали FRBR интуитивным.

Второе исследование (2010 г.) проводилось при участии 60 студентов, - по 10 человек, занимающихся компьютерными науками, дизайном, экономикой, историей, медициной и общественными науками. Мы взяли шесть графов, основываясь на моделях, анализированных в ходе первого исследования, и попросили участников выбрать их тот, который (по их ощущениям) лучше всего представлен в перечисленных библиографических сущностях и их связях. Результаты показывают, что большинством участников (33 из 60) была выбрана модель FRBR, и очевидной альтернативы этому не было.

Поскольку в ходе обоих исследований соглашение с FRBR было перевыполнено, мы обнаружили, что позиционирование оригинального выражения в FRBR ставилось под вопрос некоторыми участниками, т.к. они видели в этом важнейшее звено между произведением и остальным библиографическим массивом. Это различие не отражено в FRBR, но оно присутствует в модели FRBRoo (International Working Group, 2010). Мы будем продолжать работу, незначительно расширив поле, охваченное вторым исследованием.

В некотором роде схожее исследование (Arastoopoor, Fattahi, Pariookh, 2010) рассматривало мнения экспертов в области размещения связанных сущностей. Было выявлено, что участники понимают FRBR, но дополнительно использовали другие, более специфические критерии для определения расположения. Результаты этих исследований предполагают, что FRBR является уместной, ориентированной на пользователя концептуальной моделью – по крайней мере, для частей библиографической вселенной.

### **От теории к практике**

В последнее время появилось несколько разработок на основе FRBR, - таких как новые Международные принципы каталогизации (2009) и даже первые попытки создания Правил каталогизации на основе FRBR в виде Итальянских правил каталогизации (Trombone and Canera, 2009, Petruciani, 2009) и RDA (2010). С другой стороны, завершение еще двух концептуальных моделей семейства FRBR - FRAD (2009) и FRASAD (2010) – даёт возможность дальнейшего проникновения в различные части библиографической среды. Еще одной важной областью является разработка пространства имён, которой продолжает заниматься IFLA Namespaces Task Group.

Эти важнейшие блоки для постройки нового поколения каталогов уже на месте, но продолжается обсуждение по некоторым крупнейшим вопросам, которые по-прежнему ожидают решения (смотри, например: Yee (2005), Žumer (2006, 2007 a и b), Dickey (2007), Salaba и Zhang (2007), Zhang и Salaba (2007, 2009a, 2011a и b), Carlyle (2006), Carlyle и Fusco (2007)). В настоящем докладе мы постараемся дать отчёт о некоторых моментах.

Тринадцать лет спустя после публикации FRBR мы верим в то, что настало время для того, чтобы библиотечное сообщество поняло, что FRBR можно и нужно использовать для разработки нового подхода к каталогизации. И до сих пор кажется, что библиотекари всё ещё не осознают, каков полный потенциал сетевой библиотечной среды, и хотят по-прежнему держаться за некоторые инструменты и методики, которые утратили своё назначение в ходе автоматизации библиотек. В этом смысле инициативы, допускающие продолжение работы по современным методикам, ничем не помогут. Опыт последних лет и опыт других провайдеров говорит нам, что каталоги

для того, чтобы уцелеть, должны предоставить пользователям более эффективный способ получения и использования информации, а также – улучшенные данные. Библиотекам следует начать предлагать такие инструменты, которые обеспечат пользователей контекстом, который значительно шире доступного в каталоге сегодня. Открыв же свои инструменты и данные внешнему миру, библиотеки могли оказывать сильное влияние на развитие Семантического веба. И правда: мы преследуем общую цель – улучшенная доступность к информации и её организация. Поэтому мы должны активнее вовлекаться во взаимодействие с сообществом Семантического веба.

## Ключевые вопросы

### Ключевые вопросы

Есть несколько ключевых вопросов, на которых, по нашим ощущениям, необходимо сконцентрироваться библиотечному сообществу, не выпуская из поля своего зрения цели каталогизации.

#### 1. Разработка модели

##### а) Гармонизация (различных программ)

Хотя и FRAD, и FRSAD основываются на FRBR и следуют одному и тому же методу моделирования, разработчиками были приняты разные решения относительно моделирования, результатом чего стали решения, частично не сочетающиеся друг с другом. FRBR – концептуальная модель библиографической вселенной, определяющая все сущности и связи, но особенное внимание обращает на Группе 1. FRAD распространяет эту модель на область авторитетных данных для сущностей Группы 2 и *произведений*, FRASD занимается предметными связями.

Главные различия между FRBR и FRAD сводятся к следующему:

- К FRAD добавлены две пользовательские задачи: «Контекстуализировать» и «Подтвердить» (оправдать)
- Добавлены три «звательных» сущности: *имя (наименование)*, *идентификатор* и *контролируемая точка доступа*;
- Связь 'имеет в качестве предмета' в FRBR заменили в FRAD предметом, как атрибутом *произведения*;
- *Правила* и *Учреждение* появились в FRAD, как новые сущности.

Главные различия между FRBR и FRSAD это:

- Добавление задачи “Исследовать”;
- Введено понятие *Thema* - как суперкласс всех сущностей, которые могут быть предметами некоего *произведения*;
- В группе 3 явно не предопределены никакие сущности;
- Введено понятие *Nomen*, оно определено как отдельная сущность, а не как атрибут.

Все эти различия расхолаживают широкое внедрение FRBR, - поэтому очень важно срочно провести гармонизацию семейства FRBR. Группа по пересмотру FRBR (FRBR Review Group) уже начала этот процесс, отчасти – в связи с проектом Namespaces, но к настоящей работе еще не приступали.

С другой стороны библиотечное сообщество совместно с музейным относительно успешно разработало общую модель: FRBRoo является результатом гармонизации FRBR и CIDOC CRM и может служить основанием для разработки общих инструментов и сервисов для документирования культурного наследия.

#### б) Атрибуты и связи

Вопреки всем упомянутым выше инициативам, у нас по-прежнему очень мало полученных опытным путем свидетельств того, какие точно атрибуты и связи требуются различным группам пользователей. Модель FRBR взяла атрибуты и связи преимущественно из современных методик каталогизации: *International Standard Bibliographic Descriptions (ISBDs)*, *Guidelines for Authority and Reference Entries (GARE)*, *Guidelines for Subject Authority and Reference Entries (GSARE)* и *UNIMARC Manual*. Разработчики модели FRBR утверждают: “Идентификация и определение атрибутов для различных типов материалов могут быть расширены за счёт дальнейшего рассмотрения экспертами и изучения пользователей” (FRBR, 1998). Однако, более новые начинания, такие как RDA и более новая версия ISBD по-прежнему основываются в основном на атрибуты и связи, прописанные и перечисленные в первом исследовании FRBR.

Небольшое исследование, выполненное Leskovec (2005) подтверждает, что атрибуты и связи, записанные в современных каталогах, не всегда соответствуют потребностям пользователя. Она проанализировала запросы пользователей одной публичной библиотеки и выяснила, что большинство пользователей ищет выражения (например – любое издание конкретного произведения на конкретном языке), а иногда даже – произведение вообще. Некоторые пользователи ищут проявления (например, конкретное издание), когда им особенно интересны первое или последнее издание или когда они ищут издания с дополнительными материалами: иллюстрациями, комментариями и т.п. Однако, тогда как записи каталога подробно описывают проявления, информация о соответствующем произведении (произведениях) и выражении (выражениях) не всегда очевидна и множество важных связей и атрибутов в запись не внесено (полный текст или сокращенный, информация о продолжениях и т.д.).

## 2. FRBR-изация

Существует мощный массив старых (унаследованных) данных, которым необходимо как-то сосуществовать с данными, созданными по FRBR. Лучшее решение – FRBR-изация. FRBR-изация – это процесс извлечения из существующих библиографических записей сущностей, относящихся к FRBR. Поскольку рекаталогизация огромных объемов существующих данных не является реальным выходом, FRBR-изация жизненно важна для прозрачного совместного использования старых и новых данных.

Два года назад на Конгрессе ИФЛА мы докладывали об эксперименте с данными Национальной библиографии Словении, норвежской базы данных Bibsys и базы данных e Swedish Burk (Pisanski, Žumer, Aalberg, 2009). Мы обнаружили, что существующие данные могут быть FRBR-изованы удовлетворительным способом, но качество результатов зависит от различных важных факторов, таких как качество и

последовательность библиографических записей, местные методики каталогизации, правила каталогизации и форматы. Эти проявления непоследовательности приводили к неверной идентификации сущностей, что может затмить преимущества FRBR-изации в глазах конечного пользователя. Возможно, важнее всего то, что наш эксперимент ясно показал: FRBR-изация – это не просто, и для того, чтобы отражать отличия индивидуальных методик каталогизации, ей нужна настройка. Доклады о предшествовавших попытках проведения FRBR-изации ясно сообщают нам о том же: Hegna и Murtomaa (2002); Hickey и O'Neill (2005).

Существует много начинаний, вдохновлённых FRBR, ясно показывающих признание того, что пользователи выиграли бы, если бы информация была бы лучше структурирована. Однако, ни одна из этих попыток (кроме нескольких экспериментов небольшого масштаба) действительно полностью следовали FRBR: ведь уровень выражения часто выпускается или низводится до уровня групп выражений на основе естественного языка, чаще всего из-за сложностей с идентификацией выражений в ходе FRBR-изации. Однако, как уже упоминалось, Leskovec (2005) обнаружила, что иногда пользователи ищут частное выражение (и не всегда на конкретном языке) и определённо выиграют от такой информации.

Не существует простого способа записи всей важной информации (включая связи, которые, в конце концов должны стать центральной точкой каталогов следующего поколения) в существующих библиографических записях, которые структурированным образом используют существующий инструментарий. Это особенно верно в сложных случаях, например, - агрегаторах разного рода.

### 3. Идентификация

На прошлогоднем Конгрессе ИФЛА мы представили доклад о значимости идентификации в области библиографической вселенной, уделив особое внимание необходимости систем идентификации в многоязычных и поликультурных средах (Pisanski, Žumer, Aalberg, 2010). В наличии имеется некоторое количество идентификационных систем, однако даже самые распространённые – такие как ISBN – по-прежнему используются недостаточно.

К тому же во многих случаях присутствует разное понимание того, какую сущность должен идентифицировать данный идентификатор, что, по большей части, происходит из-за различных ожиданий, с которыми относятся к идентификаторам разные сообщества. Добавим, что многие идентификационные системы предназначены для использования с различными типами материалов (аудио, книги, сериалы и т.д.)

Технологии Семантической паутины, как и любой другой платформы, предъявляют свои требования и имеют свои недостатки (смотри Yee, 2009). Одним из этапов нашей борьбы за то, чтобы дать данным семантическое значение, является движение Связанных данных (Linked data initiative). Технически Связанные данные это данные, опубликованные в паутине так, что они машиночитаемы, их значение определено явно, они связаны с другими наборами данных и, в свою очередь, могут быть связаны с наборами внешних данных (Bizer, Heath, Berners-Lee, 2009). Одной из архитектурных предпосылок движения Связанных данных является использование URI (Унифицированных идентификаторов ресурсов) в качестве имен (наименований), т.к. они (URI) идентифицируют любые виды объектов или концептов (Berners-Lee, 2009).

Yee (2009) объясняет, как это будет работать в контексте FRBR: «[...] мы совместно будем работать над созданием URI для произведений, выражений, проявлений, лиц, организаций, мест, предметов и т.д.» Но, как продолжает автор там же: «В воздухе постоянно витает вопрос о том, какие структуры (организации и пр.) станут поддерживать совместное создание URI для сущностей Семантической паутины. Для того, чтобы на эти данные можно было положиться, нам понадобится некий способ, который даст гарантии того, что система будет находиться под контролем людей, достаточно образованных, чтобы понимать ценность чистых и точных определений сущностей, ценность самых общеизвестных предпочтительных форм (для вывода в списках многочисленных разнообразных сущностей) и стоимость предоставления доступа». Иными словами, если мы хотим выставлять наши данные в Семантической паутине, нам необходимо уникально идентифицировать каждую сущность (а также и связи).

На часть проблемы “Agent” указал проект Виртуального авторитетного файла (VIAF project <http://viaf.org>), одним из возможных решений могло бы стать использование их URI для идентификации агентов, по крайней мере для тех агентов, которые имеются в их базе данных.

Еще труднее вообразить, как приступить к идентификации всех сущностей 1 группы FRBR. Возможен подход, начинающийся с достижения идентификации на национальном уровне (задача сложная сама по себе), а затем продолжающийся созданием международного координационного органа.

Некоторые, возможно, возразят, что даже простая публикация в паутине национальных (или просто имеющих отношение к библиотекам) URI для сущностей лучше, чем ничего. Но в таком случае библиотеки утратят «власть» ради установления связей с энтузиастами или фирмами, а это и есть то, чего более всего опасается Yee (2009).

Каким бы ни был наш подход, мы должны быть уверены в том, что ничто не будет сделано за счёт простого перехода на платформу Семантической паутины. Необходимо давать определения идентификаторам вручную, и данные тоже не свяжутся сами по себе. Необходимо, чтобы связи были явными, а такое связывание лучше всех должны выполнить мы. Если подобные действия производятся автоматически или полуавтоматически, на результаты не всегда можно положиться, что уже доказала FRBR-изация, которая полагается на идентификацию. Частичные решения в этой области не приведут ни к экономии времени, ни к экономии средств.

#### 4. Презентация

При обсуждении изменений в каталогизации и каталогах часто выпадает из поля зрения одна область – аспект представления данных, связанный с созданием пользовательских интерфейсов, которые бы целенаправленно поддерживали пользовательские задачи и адекватно представляли данные. Carlyle (1997) призывал библиотечное сообщество обеспечить показ на основе связей и подчёркивал, что это изменение «следовало выполнить уже давно», но спустя 13 лет ситуация, в основном, всё та же. Классический вывод данных на показ, который всегда использовали в библиотечных каталогах, не способен показать все возможности, предоставляемые

различными новыми моделями библиографическим записям, которые выстроены на связях. Критические замечания о неспособности наших информационных систем целенаправленно размещать записи и представлять связи - в любом случае не новы, но этот вопрос никогда не выходил на первое место еще и потому что существующие методика каталогизации, правила и форматы в действительности не предоставляют хорошей основы для внедрения этих идей.

Как указывали Zhang и Salaba (2009b), модель FRBR предлагает мощный потенциал для разработки эффективных и дружелюбных пользователю системных интерфейсов и дисплеев за счет сбора записей во взаимосвязанные записи и вывода на показ иерархии и связей, основанных на модели FRBR. По сей день большинство инициатив, вызванных FRBR, пытались подступиться к размещению и представлению произведения - (выражения) – воплощения, но не концентрировались на связи произведения с произведением и других связях, обсуждаемых в этой модели.

Из-за недостаточного количества исследований техники представления FRBR-изированных записей и только начальных внедрений идей FRBR наша исследовательская группа решила изучить перспективы визуализации информации для представления FRBR-изированных записей и взаимодействия с ними, т.к. это могло бы дать возможность размещать данные, а также лучше представлять различные связи. С помощью системы-прототипа на тестирование будут поставлены три различных техники визуализации и три вида взаимодействия. Поскольку необходимо, чтобы визуализация работала одинаково хорошо и на очень больших и на очень простых наборах данных, примеры для тестирования будут построены таким образом, который представит различные возможные структурные и реляционные сложности.

Насколько нам известно, это будет первый эксперимент, в ходе которого будут протестированы возможные техники вывода данных на основе FRBR. И, будучи таковым, он не только покажет нам, есть ли у визуализации информации потенциал для поддержки процессов изучения и получения информации пользователем, но и даст другую ценную информацию. Надеемся, что в будущем в этой области будет проведено больше исследований; не только потому что они смогут предложить более глубокое понимание того, какие данные действительно необходимы для наших записей, но и потому что скоро самым важным вопросом станет приспособление и настройка пользовательских интерфейсов для нового поколения библиотечных данных. Без этого большая часть нашей работы будет утрачена для пользователей. Мы также надеемся, что платформа, которая обеспечивает поддержку внедрению расширенных библиотечных данных, может стать хорошим сигналом библиотеками и каталогизаторам: быстро и более последовательно перенимать новые методики.

### **Изучение пользователя**

Поскольку в центре внимания библиотек всего мира были и остаются пользователи, все наши решения должны основываться на наглядном опыте. Мы ощущаем, что пользователи и используемость применительно ко всем вышеупомянутым областям должны систематически и непрерывно изучаться. Наша группа выполнила свою небольшую долю исследований, были и другие несистематические попытки, - вот почему мы призываем научное сообщество присоединиться к нашим усилиям.



## **Заключение**

Мы показали некоторые результаты нашего исследования и те, вопросы, к которым надлежит обратиться, чтобы получить все возможное от последних многообещающих разработок в области каталогизации и за её пределами. Однако библиотеки быстро уходят от того времени экспериментов и неспешного удобного подхода к изменениям. Библиотечное сообщество также пытается приспособить новые идеи к старым методикам, что может оказаться бессмысленным, поскольку современные каталоги не соответствуют информационному поведению пользователей (UCL, 2008, OCLC, 2009). Пользователи считают, что каталоги сложны в использовании и недостаточно интуитивны. Вдобавок, большая их часть не желает вкладывать свое время и ресурсы в изучение деталей какой бы то ни было системы, даже если это могло бы им помочь добиться лучших результатов.

Библиотеки располагают инструментами, потенциально способными стать инструментарием мирового класса, однако их необходимо адаптировать к требованиям конечных пользователей. Цель должна быть ясна: обеспечить решение, которое пользователи однозначно посчитают благотворным. Цена застенчивости, проявленной при внедрении необходимых изменений, может оказаться даже слишком высокой. В то время, когда пользователи энергично ищут доступ к информации, контексту, содержанию, изучению и открытию, неудача библиотек в предоставлении всего перечисленного будет стоить потери читателей, а также потери лидерских позиций в информационном мире.