

ВНЕДРЕНИЕ ПИЛОТНОГО ПРОЕКТА BIBFRAME В БИБЛИОТЕКЕ ЙЕЛЬСКОГО УНИВЕРСИТЕТА

Тимоти А. Томпсон

*Библиотекарь по прикладным исследованиям метаданных
Библиотека Йельского университета
timothy.thompson@yale.edu*

Даниэль Мугабуру

*Технический ассистент по метаданным
Библиотека Йельского университета
daniel.mugaburu@yale.edu*

АБСТРАКТ

В академических библиотеках США идет процесс перехода от форматов MARC 21 к стандартам, основанным на связанных открытых данных. С 2018 по 2020 гг, библиотекари Йельского университета приняли участие в масштабном проекте Linked Data for Production и реализовали пилотный проект по внедрению стандарта BIBFRAME Библиотеки Конгресса для описания ресурсов. BIBFRAME – это гибкий стандарт, который может использоваться для специального описания с помощью специальных доменных словарей, таких как добавочной онтологии Art and Rare Materials (ARM).

Группа из 20 библиотекарей Йельского университета прошла обучение по использованию Sinopia Linked Data Editor, нового инструмента каталогизации, разработанного в Стэнфордском университете. 200 наименований было закаталогизировано до того, как проект был приостановлен из-за пандемии COVID-19. Во время относительно короткого промежутка времени, каталогизация ресурсов по стандарту Sinopia оказалась в два раза дольше по сравнению с традиционным MARC. Академические библиотеки по прежнему находятся на ранних стадиях перехода и нуждаются в дополнительном инвестировании в обучение и технологии для успешного его завершения.

Добрый день. Спасибо оргкомитету за приглашение выступить на Евразийской Конференции Академических Библиотек. Меня зовут Тимоти Томпсон и я библиотекарь по прикладным исследованиям метаданных в Библиотеке Йельского университета. Сегодня я поделюсь опытом по использованию стандарта BIBFRAME в рамках пилотного проекта, который был реализован в 2018-2020 гг. Эта презентация была подготовлена с помощью моего коллеги, Даниэля Мугабуру, члена проекта по метаданным.

- Сначала, я приведу описание проекта.
- Затем, остановимся на терминах, модели BIBFRAME и том, как она использовалась в нашем проекте.

- Наконец, я представляю Sinopia Editor, инструмент, который мы использовали для внедрения словаря BIBFRAME и создания новых библиографических описаний.

Наш проект проходил в рамках инициативы Связанные данные для производства (Linked Data for Production LD4P)). Под производством мы понимаем трансформацию в каталогизации, которая в настоящее время основана на форматах MARC 21, и адаптацию нового опыта работы со стандартами семантического веба.

Инициатива LD4P является результатом коллективной работы научных библиотек США, под руководством Стэнфордского университета, Университета Корнелл, Гарвардского университета, и Университета Айовы, где обязанности по разработке предложения и управлению проектами были возложены на Стэнфорд.

Сначала проекты LD4P финансировались грантами Фонда Эндрю Меллона. Во второй фазе, когда в проект вступил Йельский университет, инициатива получила поддержку и грант на развитие профессионального практикующего сообщества среди большого количества библиотек. В 2018 был опубликован конкурс заявок, и заинтересованные институты могли предложить свои проекты с бюджетом до \$50,000 USD. В итоге, было получено 17 заявок от библиотек США.

Главной целью второй части программы LD4P2 стала разработка нового инструмента каталогизации, Sinopia Editor, на основе принципа связанных открытых данных, а также его проектирование согласно стандартам семантического веба. Каждая из 17 библиотек использовала и тестировала редактор Sinopia в своем проекте и разрабатывала инструкции по внедрению нового подхода в каталогизации.

В нашем случае, для LD4P2 мы применили практический подход и обозначили три главные цели:

- Во-первых, нам нужно было итеративно изучить последствия перевода нашей каталогизации и рабочих процессов, а также адаптации нового подхода с использованием связанных данных. Мы хотели провести тренинги для сотрудников отдела каталогизации и предоставить им возможность получить практический опыт работы с BIBFRAME и другими стандартами связанных данных.
- Во-вторых, мы выбрали материалы для каталогизации, которые позволили бы применять не только BIBFRAME, но и другие семантические словари, особенно Онтологию материалов по искусству и редких материалов (Art and Rare Materials Ontology (ARM)), которая является ведущим стандартом, разработанным в первой фазе проекта LD4P. В первую очередь, мы работали с подборкой памфлетов и эфемеров 19 – нач. 20 вв. Некоторые из этих объектов долгое время пылились на полках архива нашего технического отдела. Т.к. они не являлись приоритетными единицами для каталогизации, это стало хорошим материалом для проекта. Мы смогли работать сколько угодно долго над каталогизацией и в будущем конвертировать записи BIBFRAME в формат MARC 21 для их индексации

в системах дискавери. Конвертация всегда будет необходимой, т.к. OPAC, которые мы сейчас используем поддерживает только форматы MARC. В настоящее время задание по конвертации по прежнему на выполнении.

- И наконец, т.к мы не могли представить нашу работу в OPAC, нам нужно было изучить возможности поиска, извлечения и визуализации BIBFRAM. Для этого мы использовали большую часть бюджета на приобретение лицензионного программного обеспечения на платформе metaphactory, которая применяется для поиска, визуализации, решений по управлению данными для связанных данных.

Наш проект основан на сотрудничестве технических отделов Йельской центральной библиотеки, Библиотеки Редких книг и манускриптов Бейнеке. Общая группа по созданию метаданных из 20 человек была разделена на три команды, ответственных за каталогизацию в Sinopia Editor. Каждый член группы должен был посвятить 20 процентов своего времени на каталогизацию в редакторе, что в общем составило один день в неделю.

В ходе выполнения проекта мы столкнулись с множеством трудностей и вынуждены были менять сроки задач. LD4P2 – комплексный проект, в котором было задействовано более 20 организаций, в связи с этим были административные задержки в управлении отдельными суб грантами. Тем не менее, самую большую сложность вызвало опоздание в получении Sinopia Editor. Изначально, программное обеспечение планировалось получить в апреле 2019, но версия 1.0 была готова к использованию только в августе.

Эти три месяца подготовки редактора мы провели разрабатывая шаблоны и рабочие процессы. Кроме того, еще три месяца с ноября 2019 по январь 2020 заняли тренинги для членов группы по метаданным. В итоге мы приступили к созданию метаданных в феврале 2020, но пандемия COVID-19 существенно сократила время на каталогизацию.

Несмотря на это, мы считаем, что проект был успешным. Мы смогли каталогизировать более 200 единиц в Sinopia Editor, что достаточно для показателя достижения целей, а также обеспечили каталогизаторам прочный практический опыт использования BIBFRAME и связанных данных на деле.

Краткий обзор модели BIBFRAME и ее применение в Йельском университете.

BIBFRAME – семантическая модель данных и словарей, разработанная Библиотекой Конгресса США. Следует отметить, что BIBFRAME не является продуктом международной организации по стандартизации; а принадлежит Библиотеке Конгресса, где существуют свои институциональные приоритеты, предпочтения и даже политические цели. Еще один интересный момент обращает внимание, тогда как форматы MARC 21 продолжают меняться из года в год и предлагают новые поля и подполя, BIBFRAME остается неизменным с версии 2.0 от 2016.

В 2011 году, когда впервые было объявлено о появлении BIBFRAME, особый акцент был сделан на то, что проект предполагает переход от MARC 21 к новому

интернет-дружественному формату. Действительно, с течением времени MARC 21 не так хорошо адаптируется под технологические достижения. Сегодня, как известно, онлайн-поисковые инструменты во многом заменяют библиотеки и их каталоги, и для множества людей оказываются первыми для поиска и извлечения информации. В свою очередь BIBFRAME был разработан со спецификациями, которые предполагают продвижение совместимости между библиографическими метаданными и всемирной сетью интернет.

Говоря о BIBFRAME, необходимо также помнить о его связи со стандартом RDA (Описание ресурса и доступ), который является международным кодом каталогизации. RDA содержит набор из 4 ключевых категорий: Труд (Work), Представление (Expression), выражение (Manifestation), Единица (Item). В BIBFRAME категория Труд сочетает характеристики RDA категорий Труд и Представление. Выражение и Экземпляр по сути эквивалентны категории Единица.

В нашем проекте, мы выбрали модель RDA для внесения признаков в категориях Труд и Представление. По техническим причинам, мы решили адаптировать новую категорию – Супертруд, чтобы передать категорию RDA Труд.

В первой фазе проекта LD4P, различные группы моделирования данных формировались для создания онтологии BIBFRAME. Эти расширения были разработаны с учетом перспективы их использования отдельными практикующими сообществами для описания особых форматов и материалов. Всего в BIBFRAME было создано 4 категории расширений: Искусство и редкие материалы, Картография, Динамические изображения, Исполненные музыкальные произведения.

Преимущество работы с BIBFRAME заключается в том, что он гибкий и позволяет вводить новые элементы из других семантических словарей, если это необходимо или просто удобно. Например, для нашего проекта, мы выбрали внедрение и разработку онтологии ARM, словаря, предназначенного расширить BIBFRAME для детального описания категории Единица.

Использование Sinopia Editor в каталогизации.

Sinopia можно найти на хостовом сайте в виде тренировочного (<https://stage.sinopia.io>) или производственного образца (<https://sinopia.io>). Пользователи должны зарегистрироваться и создать аккаунт, чтобы использовать размещенную версию инструмента. Программное обеспечение является открытым ресурсом и доступно для локальной установки (https://github.com/LD4P/sinopia_editor).

Чтобы перейти к интерфейсу каталогизации, на главной странице Sinopia нужно кликнуть на “Linked Data Editor”.

В редакторе Sinopia, каждая форма ввода требует соответствующего шаблона. Хостовый сайт предлагает шаблоны по умолчанию для различных форматов (например, монографии, сериальные издания, картографические ресурсы). В самом редакторе также есть уже включенные шаблоны, каждый из которых имеет свой идентификатор. Пользователи могут разрабатывать свои шаблоны, как это сделали мы в Йельском университете для пилотного проекта. Самые первые шаблоны

мы создали для монографий, которые имеют отдельные формы Супертруд, Труд, Экземпляр и Единица.

Разработанный рабочий процесс начинается с создания описания для категории Супертруд, которое последовательно опускается до уровня Единица. Каждое описание имеет связь с предыдущей категорией.

На уровне Супертруд мы вводим информацию о создании или источнике ресурса, такую как первичный создатель и его интеллектуальный контент, например соответствующие предметные рубрики.

Информация из формы Супертруд может быть внесена для ролей создателей объекта. Например, мы можем записать первичную роль с указанием, что это автор.

Если соответствующая персоналия представлена на внешних источниках связанных данных, которые были сконфигурированы для поиска, для этой категории в форму описания локального ресурса может быть внесен международный ресурсный идентификатор IRI (International Resource Identifier).

Если персоналия не найдена среди внешних связанных данных, имя можно внести в строку (согласно стандартным правилам форматированных имен) нажав на кнопку “Authorized Access Point.”

Форма ввода данных для уровня Представление выглядит следующим образом: в первом поле формы мы вводим гиперссылку на соответствующую категорию Супертруд, после того, как она создана.

Я остановился здесь, чтобы не превысить временной лимит презентации, но в полной демонстрации мы бы добавили информацию, соответствующую дальнейшим первичным уровням описания, а именно Экземпляра и Единицы.

В заключении, BIBFRAME представляет собой гибкий и открытый подход к каталогизации. Однако, адаптация новых инструментов, таких как Sinopia, требует значительных усилий и инвестиций в тренинги и технологии. Вначале эффективность процесса скорее всего будет низкой, до того как инструменты и работа наладятся. Например, мы заметили, что каталогизация ресурса в Sinopia почти в два раза дольше, чем в MARC 21. Для дальнейшего внедрения улучшенных и полезных метаданных, необходима коллективная работа, сотрудничество между институтами и на международном уроке.