

Авторитетные данные в Семантическом вебе

О.Н.Жлобинская

Введение

В финальном отчете группы Консорциума W3C по проблемам библиотечных связанных данных содержатся рекомендации по использованию стандартов Семантического веба и принципов связанных данных с целью сделать библиотечные ресурсы более заметными и доступными для повторного использования за пределами их первоначального библиотечного контекста. В частности, ведущим библиотекам рекомендовано в первую очередь определить приоритетные проекты, не требующие огромных затрат, но позволяющие библиотекам войти в пространство Семантического веба. По своей природе связанные данные предполагают постепенное наращивание – начав с некоего ядра, и устанавливая связи с другими наборами данных, мы можем постепенно все больше и больше расширять сеть. Наиболее естественный путь установления таких связей, как нам представляется, – это использование инструмента, привычного и широко используемого в библиотечном сообществе, – авторитетных данных. Действительно, само понятие связанных данных включает идентификацию сущностей и установление связей между ними, что, безусловно, перекликается с принципами авторитетного контроля, которые предполагают унификацию точек доступа и навигацию в среде данных. Кроме того, авторитетные данные характеризуются высоким уровнем детальности, необходимым для дифференциации объектов описания и для обеспечения точности поиска. Такая детальная информация может служить отправной точкой, а ссылки между точками доступа, зафиксированные в авторитетных файлах, обеспечат связь между различными инструментами и навигацию между ресурсами.

Массивы авторитетных данных, созданных и используемых библиотеками, насчитывают миллионы записей. Необходимо представить авторитетные данные в формате, который способен работать с внешними веб-службами, обеспечит интеграцию функциональности авторитетных данных – поиск ресурсов, просмотр ресурсов, связанных с определенной тематикой, навигацию между ресурсами, – в самые разные веб-инструменты. Для этого мы, прежде всего, должны определить модель представления данных, онтологии, которые будут использованы для того, чтобы выразить данные в RDF.

Словари и онтологии

В сообществе Семантического веба уже разработаны и широко используются многочисленные словари, построенные на основе стандартов W3C RDF, RDFS и OWL, которые, используя единую модель данных и общий подход, различаются уровнем выразительности:

- RDF – это модель описания связанных данных, использующая элементарные конструкции *Субъект – Предикат – Объект* и URI-ссылки для их обозначения;
- RDFS (RDF Schema) – семантическое расширение RDF, язык для описания словарей, или упрощенных онтологий;
- OWL – язык представления веб-онтологий; OWL основан на стандартных словарях RDF и RDFS и является их расширением. Основные элементы OWL – *классы*, представляющие собой механизм для группировки ресурсов со схожими характеристиками, и *свойства*, используемые для атрибутивной информации и для моделирования отношений между классами.

На основе RDF, RDFS и OWL строятся другие словари и онтологии, например: FOAF (Friend-of-a-Friend), SKOS (Simple Knowledge Organization System), BIBO (Bibliographic Ontology), SIOC (Semantically-Interlinked Online Communities), DOAP (Description of a Project), Music Ontology и т.д. В последние годы активно ведется работа по представлению в RDF моделей метаданных, используемых библиотеками, – Dublin Core (2005 г.), FRBR, RDA и т.д.

FOAF

FOAF (*Friend of a Friend, Друг друга*) – проект по созданию модели машиночитаемых домашних страниц и социальных сетей. Онтология FOAF основана на RDFS и OWL и предназначена для описания людей, их деятельности и отношений к другим людям и объектам.

Примитивным объектом FOAF является концепт – это может быть агент (группа, лицо, организация), документ, изображение и т.д. Характеристики концепта выражаются с помощью RDF-свойств; например, для лица – это имя, возраст, род, титул, день рождения, домашняя страница.

На рис. 1 в качестве примера приведен фрагмент описания лица (имя – Вася Пушкин, его электронный адрес, ник, имена его друзей) и соответствующее графическое представление.

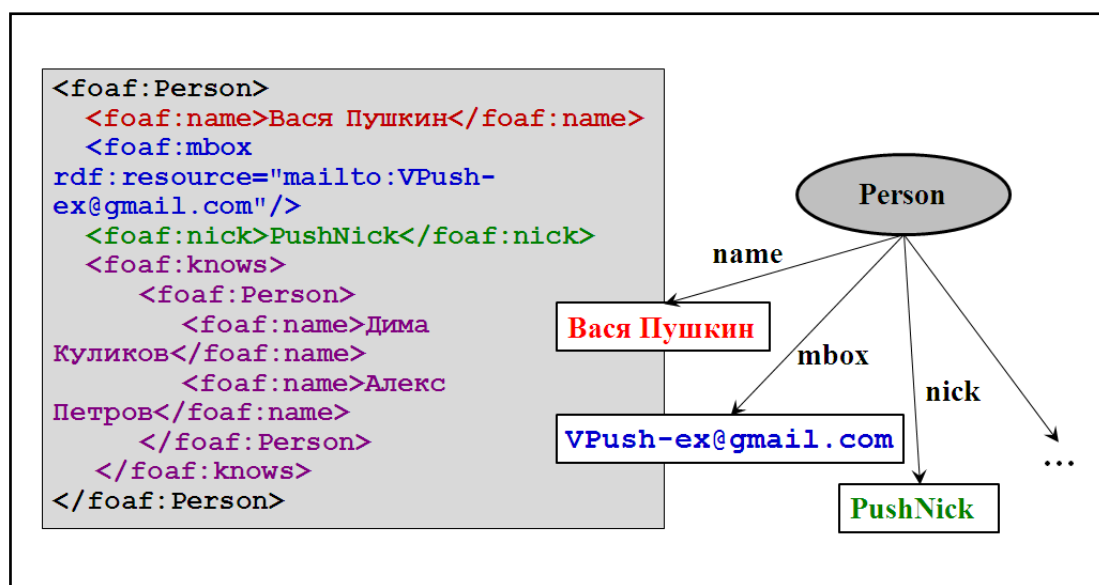


Рис. 1. Фрагмент описания лица в онтологии FOAF (RDF/XML)

SKOS

SKOS (*Simple Knowledge Organization System, Простая система организации знаний*) – модель данных для систем организации знаний – тезаурусов, классификационных схем, систем предметных рубрик; в 2009 г. SKOS опубликована в качестве рекомендации Консорциума W3C.

Основными элементами SKOS являются *концепты* и *семантические отношения*.

Концепт определяет идею, сущность, объект предметной области и является примитивным объектом SKOS. Предусмотрена возможность определения *схем концептов*, которые являются воплощением некоего словаря (можно провести аналогию с системой предметных рубрик, хотя есть важные отличия, о которых пойдет речь ниже).

Концепты представлены текстовыми метками, или *лэйблами* – это литералы, которые являются лишь атрибутами концептов, собственных атрибутов не имеют, и связи между ними также не устанавливаются.

Определены типы меток.

- **prefLabel** – *предпочтительные метки*, – считаются авторитетными метками для ресурса;
- **altLabel** – *альтернативные метки*, – могут использоваться для выражения синонимов и устранения неоднозначности меток;

- **hiddenLabel** – скрытые метки, – используются для задания информации, доступной для обработки, но скрытой от вывода, например, для указания ошибочного варианта метки.

Семантические отношения соотносят два концепта друг с другом; связи между концептами могут устанавливаться иерархические (broader/narrower; шире/уже) и неиерархические (related, ассоциирован). Возможно установление связей не только в рамках одной схемы, но между концептами, относящимися к разным схемам.

Кроме того, определены свойства для аннотирования концептов (примечания разного рода, примеры, определения).

```
<rdf:Description rdf:about="http://ex.ru/auth/ps654678">
  <skos:prefLabel>Психоллингвистика</skos:prefLabel>
  <skos:altLabel>Психология языка</skos:altLabel>
  <skos:altLabel>Язык--Психология</skos:altLabel>
  <skos:broader rdf:resource="http://ex.ru/auth/ps456421"/>
  <skos:narrower rdf:resource="http://ex.ru/auth/ps651838"/>
  <skos:exactMatch rdf:resource=
"http://id.loc.gov/authorities/subjects/sh85108432"/>
  <skos:inScheme rdf:resource="http://ex.ru/auth/subjects"/>
</rdf:Description>
```

Рис. 2. Фрагмент записи SKOS (RDF/XML)

На рис. 2 в качестве примера приведен фрагмент записи с описанием концепта: skos:prefLabel – предпочтительная метка «Психоллингвистика»; skos:altLabel – альтернативные метки «Психология языка» и «Язык--Психология»; более узкий концепт (skos:narrower), более широкий концепт (skos:broader) и отношение эквивалентности с понятием другой схемы концептов (skos:exactMatch) – представлены соответствующими URI-ссылками. Кроме того, указана схема концептов, к которой относится описываемый концепт (skos:inScheme).

При очевидном сходстве с тезаурусами и авторитетными файлами есть важное отличие – в SKOS связи устанавливаются только между концептами. В тезаурусах же и авторитетных файлах зафиксированы, как правило, связи двух типов – как между концептами (иерархические связи), так и между метками (связи типа «см.» и «ссылка от»).

SKOS-XL

Для идентификации, описания и связи лексических единиц, связанных с концептом, разработано расширение SKOS-XL (*SKOS eXtension for Labels, Расширение SKOS для меток*).

SKOS-XL позволяет приписывать меткам атрибуты и устанавливать семантические связи между метками (например, определено свойство "isTranslationOf", позволяющее установить связи между вариантами метки на разных языках).

В SKOS-XL:

- каждый концепт имеет одну или несколько предпочтительных меток на разных языках;
- каждый концепт может иметь одну или несколько альтернативных меток на каждом языке, которые считаются синонимами предпочтительной метки;
- для каждой предпочтительной метки можно указать различные литеральные формы, синонимы, статус, дату создания и другую административную информацию, относящуюся к данной конкретной метке;
- для каждой метки можно определить связи с другими метками;
- все оригинальные функции класса концептов, определенные в SKOS (например, семантические отношения между концептами, атрибуты концептов и соотнесение концептов из разных схем), применимы и к классу меток.

Таким образом, SKOS и SKOS-XL дают нам две различные модели (см. рис. 3): (1) метки определены только как свойства концептов, связи устанавливаются только между концептами (SKOS Core) и (2) метки представляют класс (SKOS + SKOS-XL), связи могут устанавливаться как между концептами, так и между метками.

Это очень коротко об основных онтологиях, которые используются для представления авторитетных связанных данных. При необходимости могут использоваться и другие словари, в нужной мере специфичные, чтобы отразить специфику данных, при условии, что они основаны на стандартных словарях и доступны для стандартных веб-инструментов.

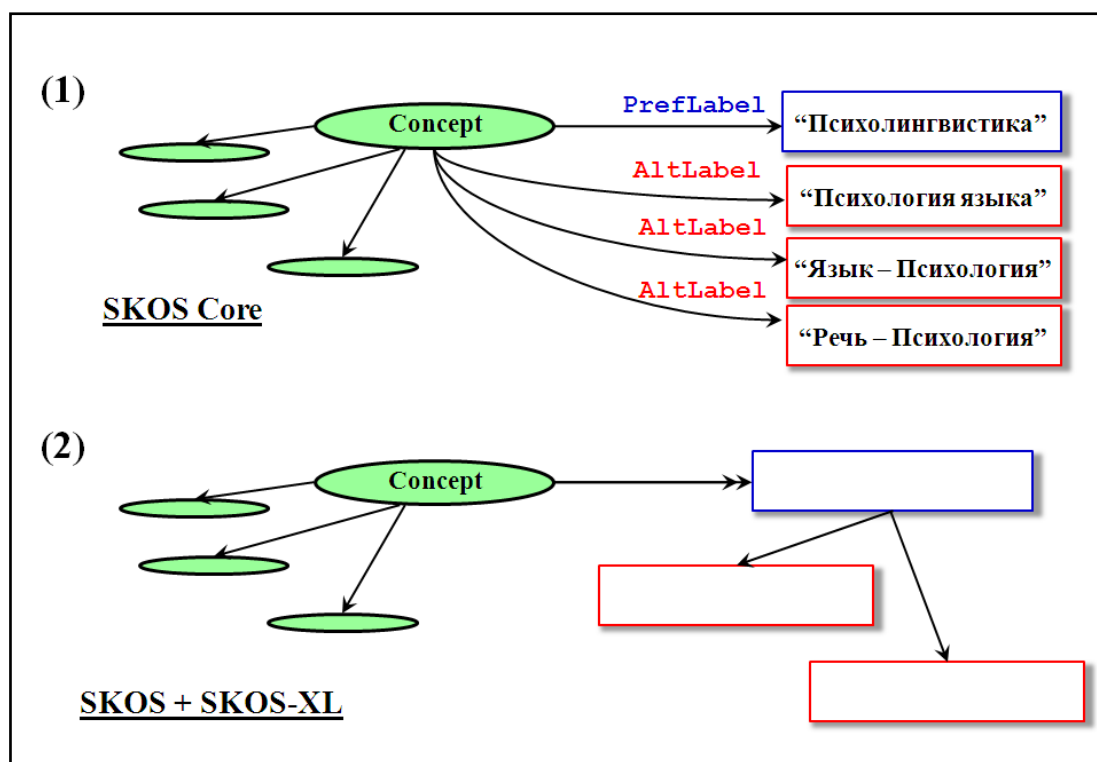


Рис. 3. Модели установления связей концептами и метками¹

Авторитетные связанные данные – практическая реализация

VIAF

VIAF (*Virtual International Authority File, Виртуальный международный авторитетный файл*) начинался как совместный проект Библиотеки Конгресса, Немецкой библиотеки, Национальной библиотеки Франции и OCLC. В 2012 г. VIAF стал службой OCLC, в настоящее время опубликован в виде связанных данных, при этом использованы онтологии: описанные выше FOAF, SKOS, а также другие – DC, RDA, VIAF и др.

На рис. 4 приведен фрагмент записи VIAF. Первая часть – указание основной сущности, описываемая в записи, и кластер имен VIAF. Далее перечислены контролируемые заголовки из различных авторитетных файлов, в данном случае – Национальной библиотеки Латвии и Библиотеки Конгресса.

¹ Использован рисунок из Žumer, M., Zeng, M. L. *Modeling knowledge organization systems and structures* (www.iskouk.org/conf2011/ppt/zumer.pdf).

```

<rdf:Description rdf:about="http://viaf.org/viaf/121053498">
  <rdf:type rdf:resource="http://xmlns.com/foaf/0.1/Person" />
  <rdf:type
rdf:resource="http://rdvocab.info/uri/schema/FRBEntitiesRDA/Person" />
  <foaf:name>Skvortsov, V. V. (Vladimir Viktorovich), 1931-</foaf:name>
  <foaf:name>Скворцов, В. В. (Владимир Викторович), 1931-</foaf:name>
  <rdaGr2:dateOfBirth>1931-01-20</rdaGr2:dateOfBirth>
</rdf:Description>
<skos:Concept rdf:about="http://viaf.org/viaf/sourceID/LNB%7CLNC10-
000044452#skos:Concept">
  <skos:inScheme rdf:resource="http://viaf.org/authorityScheme/LNB" />
  <skos:prefLabel>Скворцов, В. В. (Владимир Викторович), 1931-
</skos:prefLabel>
  <foaf:focus rdf:resource="http://viaf.org/viaf/121053498" />
</skos:Concept>
<skos:Concept
rdf:about="http://viaf.org/viaf/sourceID/LC%7Cn+2010027677#skos:Concept">
  <skos:inScheme rdf:resource="http://viaf.org/authorityScheme/LC" />
  <skos:prefLabel>Skvortsov, V. V. (Vladimir Viktorovich), 1931-
</skos:prefLabel>
  <skos:altLabel>Скворцов, В. В. (Владимир Викторович), 1931</skos:altLabel>
  <skos:exactMatch
rdf:resource="http://id.loc.gov/authorities/names/n2010027677" />
  <foaf:focus rdf:resource="http://viaf.org/viaf/121053498" />
</skos:Concept>

```

Рис. 4. Фрагмент записи VIAF (RDF/XML)

FAST

FAST (*Faceted Application of Subject Terminology*, *Фасетное применение предметной терминологии*) – фасетная система предметных рубрик, созданная на основе LCSH (*Предметных рубрик Библиотеки Конгресса*). FAST – совместный проект OCLC и Библиотеки Конгресса (1998), цель которого – адаптировать LCSH, упростив синтаксис, и при этом сохранив богатый словарь системы предметных рубрик. Словарь FAST включает 8 фасетов (Personal names, Corporate names, Geographic names, Events, Titles, Time periods, Topics, и Form/Genre), – всего около 1,7 млн. заголовков.

Для публикации FAST в виде связанных данных использованы онтологии SKOS, DC, DCT, NACO, Schema и др.

```

<skos:Concept rdf:about="1896555">
  < dct:identifier>1896555</dct:identifier>
  <skos:inScheme rdf:resource="ontology/1.0/#fast"/>
  <skos:inScheme rdf:resource="ontology/1.0/#facet-
Topical"/>
  <skos:prefLabel>Linked data</skos:prefLabel>
  <skos:altLabel>Linked open data</skos:altLabel>
  <skos:altLabel>LOD (Linked data)</skos:altLabel>
  <skos:related>
    <skos:Concept rdf:about="1017519">
      <rdfs:label>Metadata</rdfs:label>
    </skos:Concept>
  </skos:related>
  <skos:related>
    <skos:Concept rdf:about="1112076">
      <rdfs:label>Semantic Web</rdfs:label>
    </skos:Concept>
  </skos:related>

```

Рис. 5. Фрагмент записи FAST (RDF/XML)

В примере на рис. 5 приведена предпочтительная метка (`skos:prefLabel`) “Linked data”, альтернативные метки (`skos:altLabel`) “Linked open data” и “LOD (Linked data)”, ассоциативно связанные метки (`skos:related`) “Metadata” и “Semantic Web”.

Цепочки FAST, состоящие из нескольких терминов (подзаголовков), записываются в виде единой строки (см. в примере ниже “Psycholinguistics--Data processing”). Могут устанавливаться связи с концептами из других схем (`skos:relatedMatch`), в данном случае установлены ассоциативные связи с концептами, обозначенными предметными рубриками Библиотеки Конгресса, например:

```
<rdf:Description rdf:about="1081325">
  < dct:identifier>1081325</dct:identifier>
  < skos:inScheme rdf:resource="ontology/1.0/#fast"/>
  < rdf:type rdf:resource="http://schema.org/Topic"/>
  < skos:inScheme rdf:resource="ontology/1.0/#facet-Topical"/>
  < skos:prefLabel>Psycholinguistics--Data processing</skos:prefLabel>
  < schema:name>Psycholinguistics--Data processing</schema:name>
  < skos:relatedMatch>
    < skos:Concept
rdf:about="http://id.loc.gov/authorities/subjects/sh85108432">
      < rdfs:label>Psycholinguistics</rdfs:label>
    </skos:Concept>
  </skos:relatedMatch>
  < skos:relatedMatch>
    < skos:Concept
rdf:about="http://id.loc.gov/authorities/subjects/sh99005487">
      < rdfs:label>--Data processing</rdfs:label>
    </skos:Concept>
  </skos:relatedMatch>
</rdf:Description>
```

LCSH

Предметные рубрики Библиотеки Конгресса опубликованы в виде связанных данных в рамках инициативы «Linked Data Service», наряду с авторитетным файлом имен, классификационными индексами Библиотеки Конгресса, Тезаурусом для графических материалов, списками кодов для использования в MARC-записях, перечнями кодов RDA, другими словарями.

Для пользователя предлагается несколько вариантов представления и выгрузки предметных рубрик.

SKOS/RDF

При разработке модели представления системы предметных рубрик Библиотеки Конгресса авторы исходили из того, что в основе системы предметных рубрик, как и SKOS, лежит концепт. MARC-записи соответствуют абстрактным сущностям,

обозначенным заголовками – цепочками символов, и включают разные формы заголовков, семантические связи и примечания разного рода.

URI концепта строится из контрольного номера Библиотеки Конгресса по модели `http://lcsb.info/{lccn}#concept`.

При конвертировании принятый заголовок записывается как предпочтительная метка (`skos:prefLabel`), варианты заголовки – как альтернативные метки (`skos:altLabel`).

Семантические связи между принятыми заголовками, которые в MARC-записях отражаются в полях блока 5XX, в SKOS выражаются с помощью свойств `skos:narrower`, `skos:broader`, `skos:related`. Примечания также конвертируются в соответствующие свойства SKOS (`skos:note`, `skos:editorialNote`, `skos:definition`, `skos:scopeNote`, `skos:changeNote`, `skos:historyNote`).

```
-<rdf:RDF>
- <rdf:Description rdf:about="http://id.loc.gov/authorities/subjects/sh85108433">
  <rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core#Concept"/>
  <skos:prefLabel xml:lang="en">Psycholinguistics--Ability testing</skos:prefLabel>
- <skosxl:altLabel>
  - <rdf:Description>
    <rdf:type rdf:resource="http://www.w3.org/2008/05/skos-xl#Label"/>
    <skosxl:literalForm xml:lang="en">Psycholinguistic testing</skosxl:literalForm>
  </rdf:Description>
</skosxl:altLabel>
<skos:broader rdf:resource="http://id.loc.gov/authorities/subjects/sh85108448"/>
<skos:narrower rdf:resource="http://id.loc.gov/authorities/subjects/sh2003000601"/>
<skos:narrower rdf:resource="http://id.loc.gov/authorities/subjects/sh85064323"/>
<skos:narrower rdf:resource="http://id.loc.gov/authorities/subjects/sh85113741"/>
<skos:narrower rdf:resource="http://id.loc.gov/authorities/subjects/sh85011674"/>
<skos:inScheme rdf:resource="http://id.loc.gov/authorities/subjects"/>
<skos:altLabel xml:lang="en">Psycholinguistic testing</skos:altLabel>
+ <skos:changeNote></skos:changeNote>
+ <skos:changeNote></skos:changeNote>
</rdf:Description>
</rdf:RDF>
```

Рис. 6. Фрагмент записи LCSH (RDF/XML) (“Psycholinguistics--Ability testing”)

Запись на рис. 6 описывает концепт, представленный составной предметной рубрикой. Указаны предпочтительная метка (принятый заголовок “Psycholinguistics--Ability testing”), альтернативный заголовок “Psycholinguistic testing”, а также связи с более широкими и более узкими понятиями (напомним, что в SKOS связи устанавливаются не между метками, а между концептами).

Однако при таком представлении значительная часть информации теряется: в частности, нет возможности указать тип концепта (тема, географический концепт и

т.д.), информацию об элементах предметной рубрики, часть административной информации об авторитетной записи.

MADS/RDF

Для более детального представления авторитетных данных используется MADS/RDF (*Metadata Authority Description Schema in RDF, Схема описания авторитетных метаданных в RDF*) – представление модели данных MADS в виде OWL-онтологии.²

В то время как SKOS/RDF имеет очень широкую область применения, MADS/RDF разработана специально для поддержки авторитетных данных, используемых в библиотечных системах.

В частности, в отличие от SKOS/RDF, где предметные рубрики Библиотеки Конгресса представлены в «плоском виде», т.е. элементы предкоординированных цепочек объединены в одну строку (см. «Psycholinguistics--Ability testing» в примере выше), MADS/RDF позволяет определить как составную предметную рубрику в целом, так и ее отдельные элементы, и обеспечивает поддержку предкоординированных цепочек, которые составляют существенную часть предметных рубрик. Вопрос о необходимости предкоординации обсуждается в библиотечном сообществе уже много лет. Системы предметных (в частности, и LCSH) создавались для карточных каталогов, и в то время преимущества предкоординации практически не вызывали сомнений. С введением компьютеризированных процессов каталогизации и поиска в ней больше нет необходимости. Сейчас противники предкоординации получили новые аргументы – они отмечают, что координация по сути противоречит понятиям, используемым в мире связанных данных – использование отдельных кирпичиков-элементов, из которых строятся предметные рубрики, а не предкоординированных цепочек, обеспечит большую гибкость. Тем не менее, единого мнения так и нет.³

Кроме предкоординации, в MADS/RDF поддерживаются различные типы авторитетных данных – имя, тематический заголовок, географическое название и т.д., принятые и варианты метки, устаревшие формы, отдельные части меток (например, несортируемая часть заглавия в АФ заглавий).

² MADS (*Metadata Authority Description Schema, Схема описания авторитетных метаданных*) – XML-схема, разработанная Библиотекой Конгресса США для представления набора элементов, которые могут быть использованы для обеспечения используемых в библиотечном сообществе метаданных об агентах (людях, организациях), событиях, тематических терминах и т.д.

³ При обсуждении развития SKOS вопрос о поддержке предкоординации также ставился, но решение было отложено на неопределенное время.

```

- <rdf:RDF>
- <madsrdf:ComplexSubject rdf:about="http://id.loc.gov/authorities/subjects/sh85108433">
  <rdf:type rdf:resource="http://www.loc.gov/mads/rdf/v1#Authority"/>
  <madsrdf:authoritativeLabel xml:lang="en">Psycholinguistics--Ability
  testing</madsrdf:authoritativeLabel>
- <madsrdf:componentList rdf:parseType="Collection">
  <madsrdf:Topic rdf:about="http://id.loc.gov/authorities/subjects/sh85108432">
  - <madsrdf:Topic>
    <rdf:type rdf:resource="http://www.loc.gov/mads/rdf/v1#Authority"/>
    <madsrdf:authoritativeLabel xml:lang="en">Ability testing</madsrdf:authoritativeLabel>
  - <madsrdf:elementList rdf:parseType="Collection">
    - <madsrdf:TopicElement>
      <madsrdf:elementValue xml:lang="en">Ability testing</madsrdf:elementValue>
    </madsrdf:TopicElement>
    </madsrdf:elementList>
  </madsrdf:Topic>
  </madsrdf:componentList>
</madsrdf:ComplexSubject>
</rdf:RDF>

```

Рис. 6. Фрагмент записи LCSH (MADS – RDF/XML) (“Psycholinguistics--Ability testing”)

В приведенном на рис. 6 фрагменте записи MADS/RDF представлена предметная рубрика «Psycholinguistics--Ability testing». Помимо «плоской» предкоординированной цепочки (`madsrdf:authoritativeLabel` внутри тега `madsrdf:ComplexSubject`), запись включает составные части предметной рубрики (`madsrdf:componentList`), в данном случае – заголовок предметной рубрики, представленный URI <http://id.loc.gov/authorities/subjects/sh85108432>, и подзаголовок «Ability testing» (`madsrdf:topicElement`). И заголовок, и подзаголовок относятся к классу `madsrdf:Topic`.

Еще одна важная особенность MADS/RDF – это возможность провести различие между сущностью реального мира и концептуальным представлением этой сущности (авторитетными данными), указав свойства как одного, так и другого. В MADS/RDF определен класс `madsrdf:RWO` – абстрактная сущность, идентифицирует объект реального мира, и свойства `madsrdf:Authority` и `madsrdf:identifiesRWO`, которые позволяют связать концепт с объектом реального мира, который он представляет.

```

- <rdf:RDF>
- <madsrdf:PersonalName rdf:about="http://id.loc.gov/authorities/names/n79139258">
  <rdf:type rdf:resource="http://www.loc.gov/mads/rdf/v1#Authority"/>
  <madsrdf:authoritativeLabel xml:lang="en">Bradbury, Ray, 1920-2012</madsrdf:authoritativeLabel>
- <madsrdf:elementList rdf:parseType="Collection">
  - <madsrdf:FullNameElement>
    <madsrdf:elementValue xml:lang="en">Bradbury, Ray,</madsrdf:elementValue>
  </madsrdf:FullNameElement>
  - <madsrdf:DateNameElement>
    <madsrdf:elementValue xml:lang="en">1920-2012</madsrdf:elementValue>
  </madsrdf:DateNameElement>
</madsrdf:elementList>
  <madsrdf:classification>PS3503.R167</madsrdf:classification>
- <madsrdf:identifiesRWO>
  - <madsrdf:RWO>
    <rdf:type rdf:resource="http://xmlns.com/foaf/0.1/Person"/>
    <madsrdf:birthdate>19200822</madsrdf:birthdate>
    <madsrdf:deathdate>20120605</madsrdf:deathdate>
    <madsrdf:fieldOfActivity>Science fiction</madsrdf:fieldOfActivity>
    <madsrdf:fieldOfActivity>Literature</madsrdf:fieldOfActivity>
  </madsrdf:RWO>
</madsrdf:identifiesRWO>
  </madsrdf:PersonalName>
</rdf:RDF>

```

Рис. 7. Фрагмент записи MADS – RDF/XML – концепт и объект реального мира, который он представляет

Авторитетные данные как концептуальное представление объектов реального мира

Вопрос о необходимости различения объекта реального мира и его концептуального представления, в частности, обсуждался группой Консорциума W3C по проблемам библиотечных связанных данных.

Авторитетные данные часто представляют собой концептуальное представление объектов реального мира (людей, организаций, мест, событий и т.д.), которое может сопровождаться собственной мета-информацией, например, информацией о дате создания или специфическими примечаниями. С другой стороны, объекты реального мира, которым принадлежат концептуальные представления, обладают собственными атрибутами, которые описывают только их характеристики, например, дата рождения для людей, дата исторического события, широта и долгота географического объекта. Поэтому имеет смысл различать сущность реального мира и концептуальное представление этой сущности (Cluster Authority Data. http://www.w3.org/2005/Incubator/lld/wiki/Cluster_Authority_data).

Далее: различные агентства неизбежно создают свои собственные концептуальные представления (с различными URI) одних и тех же объектов реального мира и снабжают эти представления своей мета-информацией. Связывание этих концептуальных представлений с помощью утверждения `owl:sameAs` некорректно, так как в этом случае атрибуты всех метаданных были бы отнесены ко всем URI. Тем

не менее, связь между ними существует, и она должна быть отражена другими средствами.

Для решения обеих указанных выше проблем – различения и связывания концептуальных представлений и объектов реального мира Группа Консорциума W3C по проблемам библиотечных связанных данных предлагает следующую схему:

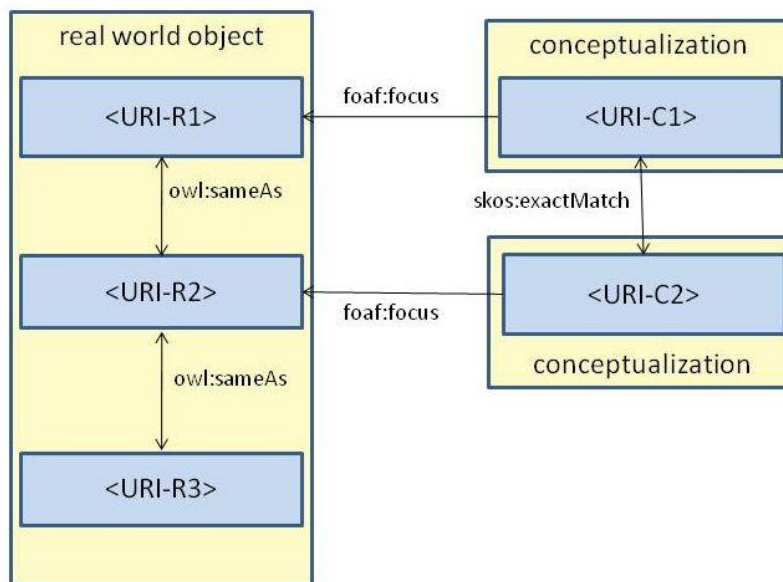


Рис. 7. Связь объектов реального мира и их концептуальных представлений

- Для связи URI, относящихся к одному и тому же объекту реального мира, используется свойство `owl:sameAs`:

`<URI-R1> owl:sameAs <URI-R2>` и `<URI-R2> owl:sameAs <URI-R3>`

- Концепт (а точнее, URI концепта) в авторитетном файле связан с URI соответствующего объекта реального мира с помощью свойства `foaf:focus`:

`<URI-C1> foaf:focus <URI-R1>` и `<URI-C2> foaf:focus <URI-R2>`

- Для выражения отношения эквивалентности между концептами, относящимися к разным схемам концептов, используется свойство `skos:exactMatch`, которое указывает на то, что эти концепты являются взаимозаменяемыми при информационном поиске:

`<URI-C1> skos:exactMatch <URI-C2>`

Таким образом, для связывания авторитетных данных используется тот факт, что они относятся к одному и тому же объекту реального мира. На рис. 8 представлен пример реализации такого подхода в Национальной библиотеке Швеции).

```

- <rdf:Description rdf:about="http://libris.kb.se/resource/auth/207420#concept">
  <skos:exactMatch rdf:resource="http://viaf.org/viaf/102333412/#skos:Concept"/>
  <skos:altLabel>Dzjejn Osten</skos:altLabel>
  <skos:altLabel>Osten, Dzjejn, 1775-1817</skos:altLabel>
  <skos:altLabel>Jane Austen</skos:altLabel>
  <skos:altLabel>Austen, Jane, 1775-1817</skos:altLabel>
  <foaf:focus rdf:resource="http://libris.kb.se/resource/auth/207420"/>
  <rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core#Concept"/>
</rdf:Description>
- <rdf:Description rdf:about="http://libris.kb.se/resource/auth/207420">
  <rda:dateOfDeath>1817</rda:dateOfDeath>
  <owl:sameAs rdf:resource="http://id.loc.gov/authorities/names/352094"/>
  <foaf:name>Dzjejn Osten</foaf:name>
  <dbpedia:birthYear>1775</dbpedia:birthYear>
  <dbpedia:deathYear>1817</dbpedia:deathYear>
  <rdf:type rdf:resource="http://xmlns.com/foaf/0.1/Person"/>
  <foaf:name>Jane Austen</foaf:name>
  <foaf:name>Austen, Jane, 1775-1817</foaf:name>
  <rdfs:isDefinedBy rdf:resource="http://data.libris.kb.se/open/auth/207420.rdf"/>
  <owl:sameAs rdf:resource="http://viaf.org/viaf/102333412"/>
  <rdf:seeAlso rdf:resource="http://en.wikipedia.org/wiki/Jane_Austen"/>
  <owl:sameAs rdf:resource="http://dbpedia.org/resource/Jane_Austen"/>
  <foaf:name>Osten, Dzjejn, 1775-1817</foaf:name>
  <rda:dateOfBirth>1775</rda:dateOfBirth>
  <libris:key>Austen, Jane, 1775-1817</libris:key>
</rdf:Description>

```

Рис. 8. Фрагмент записи LIBRIS Authority (Национальная библиотека Швеции)

Набор словарей и онтологий, которые используют для публикации связанных данных, может быть различным. Вероятно, и нет единственно верного пути, который был бы идеален для всех. Единственное условие, которое требует неукоснительного соблюдения – при разработке модели необходимо максимально использовать уже существующие словари, а в случае, если их окажется недостаточно – при создании своего словаря обеспечить его полную совместимость со словарями, построенными на основе стандартов W3C RDF, RDFS и OWL.