

# Отображение модели данных RDF в каноническую модель предметных посредников

© Н. А. Скворцов  
Институт проблем информатики РАН  
Москва  
nskv@ipi.ac.ru

## Аннотация

Модель данных RDF предназначена для описания ресурсов произвольного вида в открытой информационной среде и их идентификации в ней. При решении задач над множественными неоднородными информационными ресурсами возникает необходимость использовать данные, представленные в данной модели. В статье рассмотрены подходы к отображению модели данных RDF и сопутствующих ей языков RDF-Schema и SPARQL в объектную модель языка СИНТЕЗ, используемого в качестве унифицирующей информационной модели при интеграции множественных неоднородных информационных ресурсов в информационные системы, создаваемые на основе технологии предметных посредников. Предложен подход, учитывающий выявление семантики данных RDF и схемы данных в случае её присутствия.

*Работа выполнена при поддержке РФФИ (гранты 11-07-00402-а, 13-07-00579-а) и Президиума РАН (программа 16П, проект 4.2).*

## 1 Введение

Исследования, посвящённые решению научных задач над множественными неоднородными информационными ресурсами, включают разработку предметных посредников [14]. Предметный посредник определяется: канонической унифицирующей моделью данных, необходимой для однородного представления информации и манипулирования ею; спецификациями онтологии, описывающей предметную область посредством понятий, и концептуальной схемы, определяющей

структуру и поведение информационных объектов предметной области в терминах канонической модели; набором информационных ресурсов, интегрированных в предметную область.

Для интеграции информационных ресурсов в спецификацию предметной области посредника модели данных, в которых представлены интерфейсы информационных ресурсов, отображаются в унифицирующую каноническую модель посредника, а после отображения моделей данных схемы данных ресурсов в канонической модели отображаются в спецификации предметной области. Отображение схем представляется в виде взглядов над концептуальной схемой предметной области посредника.

Каноническая модель посредника представляет собой ядро, в качестве которого используется расширяемая модель данных, определяемая языком СИНТЕЗ [8], и набор расширений в терминах ядра, соответствующих моделям интегрируемых в посредник информационных ресурсов. Разработка канонической модели посредника состоит в унификации моделей, при которой модели информационных ресурсов отображаются в ядро канонической модели, и формируют новые расширения, либо отображаются в существующие [7].

Неоднородность информационных ресурсов, которые могут быть интегрированы в предметные посредники, понимается широко. Она включает как неоднородность на уровне схем, так и практически произвольные модели данных. Помимо методов отображения в каноническую модель традиционных моделей данных, таких как объектная и реляционная модели, в последнее время были рассмотрены принципы отображения в каноническую модель различных перспективных моделей данных, включая сервисы, онтологические модели [18], модели многомерных массивов данных [16], модели баз данных NoSQL [15]. Рассматриваются также графовые модели [17].

Ещё одной широко используемой моделью данных является модель RDF [rdf]. Она используется для описания ресурсов произвольного вида в открытой информационной среде, а также в качестве базовой модели в проектах

---

Труды 15-й Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» — RCDL-2013, Ярославль, Россия, 14-17 октября 2013 г.

Семантического Веба. В основе модели лежат триплеты «субъект-предикат-объект», посредством которых описываются ресурсы.

RDF является разновидностью графовой модели данных, в которой субъекты и объекты являются узлами, а предикаты – направленными именованными рёбрами. Однако относительно графов модель RDF привносит свою специфику. Во-первых, посредством использования URI она определяет глобальную идентификацию ресурсов в открытом информационном пространстве. Во-вторых, модель является расширяемой. За рёбрами RDF-графов может скрываться определённая семантика в зависимости от используемых словарей. В частности, RDF-модель используется в качестве базиса расширяющих её моделей, в которых предопределены предикаты, несущие конкретную семантику элементов моделей. В-третьих, в модели RDF предусмотрены правила вывода неявных связей, которые также определяют их семантику элементов модели. В зависимости от словарей, определяющих семантику описаний, набор правил вывода также может расширяться. Все эти особенности отражаются на подходах к отображению данной модели в каноническую модель предметных посредников. Таким образом, модель данных RDF заслуживает отдельного рассмотрения.

С моделью RDF неразрывно связаны модель RDF-Schema [2], предоставляющая средства определения схем данных для RDF-описаний, а также язык запросов SPARQL [3], основанный на использовании триплетов в запросах. Фактически они вместе определяют обогащённую модель данных, включающую языки определения данных и манипулирования данными, для описаний в модели RDF. Помимо этого, модель RDF является базисом для других моделей данных. В частности, язык OWL [1] со определёнными семантикой и набором правил вывода, в одном из своих синтаксисов основывается на RDF, и данные, представленные в модели OWL, обретают вид RDF-ресурсов.

Данные RDF могут быть распределены в Вебе и представляться в открытой среде в виде документов, сгенерированных в соответствии с одним из синтаксисов, представляемых для модели RDF. Для поиска данных служат специализированные поисковые механизмы, обращение к которым возможно с помощью языка SPARQL [3]. С другой стороны, для хранения и манипулирования RDF-описаниями разработан ряд специализированных хранилищ. Большинство из них использовали реляционные отношения для представления графов и триплетов в них. Сегодня RDF-хранилища используют колоночную организацию хранения триплетов и горизонтальное масштабирование [6]. Помимо этого, множество проектов в Вебе предоставляют API-интерфейс к своим данным в виде RDF в соответствии с определённой схемой и точек доступа SPARQL.

Таким образом, информационными ресурсами с данными RDF, интегрируемыми в предметные посредники, могут становиться поисковые механизмы с точками доступа SPARQL общего назначения, RDF-хранилища или Веб-проекты с API, представляющим данные в виде RDF.

Правила вывода [2] над RDF-описаниями могут прорабатывать в разное время в зависимости от системы и соображений эффективности. Обычно они применяются при обновлении данных для обеспечения эффективности запросов, однако они могут вызываться и при выполнении запросов, если обеспечение скорости обновления данных важнее скорости выполнения запросов. Некоторые RDF-хранилища могут не поддерживать вывод, отключать его, обеспечивать вывод только подмножества правил. Состав правил может быть и расширен, например, за счёт использования производных моделей, таких как OWL [1]. При интеграции информационных ресурсов, предоставляющих данные в модели RDF и при отображении модели эти особенности должны быть учтены.

Дальнейшее изложение посвящено описанию отображаемой и целевой моделей (раздел 2), опыту отображения RDF в других исследованиях (раздел 3), выбору подхода для отображения модели RDF в спецификации в унифицирующей модели, используемой в среде предметных посредников, и отображению языков RDF, RDF Schema и SPARQL в язык СИНТЕЗ (раздел 4).

## 2 Связанные работы

Отображение и трансляция данных между моделью RDF и другими моделями данных не редкость. Большинство данных хранится и обрабатывается в моделях, отличных от RDF, а трансляция их в RDF необходима для обмена информацией между проектами, между различными представлениями данных, для публикации данных в семантическом вебе в таких проектах как Linked Open Data [12] и так далее. В основном, разрабатываемые отображения имеют следующие направления.

Большинство инструментов для работы с RDF используют исключительно триплеты в интерфейсе для всех предикатов, включая определённые в самих моделях RDF и RDF Schema. При этом сами инструменты могут быть предназначены для использования в других моделях данных, например, с объектно-ориентированными языками программирования. Отображение RDF в другие модели в них производится без ухода от триплетной формы [4], а основные преобразования из моделей, используемых в инструменте, в RDF приходится совершать на уровне прикладных программ.

В реляционных хранилищах триплетов интересно рассмотреть подходы к их хранению. Первый из них заключается в хранении триплетов в единственном отношении с тремя основными

атрибутами, соответствующими субъекту, предикату и объекту. Помимо этого в нём же могут храниться идентификаторы графов и другие необходимые элементы. Чтобы не хранить все данные в одном большом отношении, его разбивают по некоторому принципу, например, по принадлежности субъектов определённым классам [11]. Но для этого над RDF-данным должна быть определена некоторая схема. Другой способ разбиения отношения – по предикату [virt], при котором в отношении хранятся субъекты и объекты, связанные посредством одного и того же предиката. Предлагаемые представления продолжают поступать [13], мотивируемые поиском эффективного хранения и доступа к RDF-графам.

Отображения, учитывающие семантику данных, обычно разрабатывается для обратной задачи: трансляции данных из реляционной или объектной модели в RDF [10]. В этом случае обратная трансляция должна восстановить исходные спецификации.

Работы, связанные с отображением языка RDF с учётом семантики данных в высокоуровневые модели данных, разрабатываются для работы с RDF-данными из языков программирования [9] или языков декларативных запросов [5].

### 3 Модели данных RDF и СИНТЕЗ

Модель данных RDF включает в себя спецификации триплетов «субъект-предикат-объект», из которых строится направленный помеченный граф с субъектами и объектами в узлах и предикатами в дугах. Семантика предикатов и ресурсов определяется пространствами имён, соответствующими словарям. С ресурсами, фигурирующими в субъектах и объектах, связаны глобальные идентификаторы URI [uri]. Допускаются безымянные узлы, не имеющие идентификации. Также предусмотрено использование в качестве объектов в триплетах примитивных типов данных, определённых в XML-Schema [xmls], а также контейнеры (bag, seq, alt) и коллекции (list). Модель RDF определяет несколько синтаксисов сериализации описаний. Определяется графический синтаксис, два подхода к представлению триплетов, XML-синтаксис.

RDF Schema [2] привносит спецификации множеств ресурсов для определения схем данных RDF. Определены классы ресурсов, принадлежность к которым определяется предикатом type, и свойства с указанием классов области определения и области значений. Введены отношения обобщения/специализации для классов и для свойств. Определён список простых правил вывода [2], таких как транзитивность и рефлексивность отношений обобщения/специализации, отнесение ресурсов к классу в соответствии с областью определения или значения свойства, правила ошестствления и другие. Они должны выполняться над спецификациями для выявления новых связей.

Язык запросов SPARQL [3] включает в себя средства задания запросов в виде конъюнкции триплетов, в которых некоторые элементы могут быть переменными. Помимо триплетной записи предоставляются возможности задания объединения, фильтров с помощью выражений, отметки необязательных элементов, сортировки, исключения дубликатов, предикативных запросов и другие. В заголовке запроса указываются используемые пространства имён. Результаты возвращаются в виде списка кортежей в соответствии с указанной структурой в виде XML, JSON или в виде триплетов.

Целевой моделью для отображения этих языков, связанных с моделью RDF, является язык СИНТЕЗ [8], выполняющий роль ядра для формирования канонической модели посредников. Это язык спецификации информационных ресурсов, который включает в качестве синтаксической основы язык фреймов, построенную над ним объектную модель и средства выражения логических формул.

Фреймы имеют идентификатор и набор слотов, каждый из которых может иметь набор значений слота. Язык фреймов позволяет специфицировать метафреймы, метаслоты и метазначения, которые сами определяются как фреймы. Средствами языка фреймов выражаются произвольные виды информации, в том числе, слабоструктурированные.

Объектная модель рассматривает фреймы как типизированные значения. Фрейм может отражать состояние объекта определённого типа, в этом случае, его структура должна соответствовать спецификации некоторого абстрактного типа данных. Помимо абстрактных типов данных, определяющих структуру и поведение объектов, вводятся также классы как множества однотипных объектов. Фрейм может становиться экземпляром класса, если соответствует типу экземпляров этого класса. Особым видом класса является метакласс ассоциаций, представляющий собой множества ассоциаций, между объектами в соответствии со спецификациями атрибутов типов.

Язык формул в язык СИНТЕЗ используется для предикативных спецификаций, задающих ограничения целостности в типах, для определения правил в логических программах и для задания запросов над спецификациями задач или информационных ресурсов.

Приведёнными средствами языка СИНТЕЗ мы выражаем спецификации, отображённые из модели RDF.

### 4 Подход к отображению модели RDF с сохранением семантики данных

Тривиальным отображением модели данных RDF в каноническую модель могло бы стать отображение отношений, хранящих триплеты, в абстрактный тип с атрибутами, соответствующими субъектам, предикатам и объектам. Однако такой подход не приемлем в среде предметных

посредников. Для них важна семантическая интеграция информационных ресурсов, при которой все обнаруживаемые связи данных с семантикой предметной области должны отображаться в понятии предметной области посредника и в соответствующие структуры концептуальной схемы посредника. Поэтому выявленную семантику предметной области следует применить для создания соответствующих типов в спецификациях информационных ресурсов, интегрируемых в посредник.

#### 4.1 Отображение модели RDF во фреймовую модель

Отображать данные в базовой модели RDF в каноническую модель необходимо на уровне фреймов языка СИНТЕЗ. Фреймы, как и графы RDF, призваны описывать ресурсы произвольной природы. И при отображении RDF триплеты с общим субъектом как описывающие один и тот же ресурс должны ассоциироваться с определённым фреймом. Далее в примерах RDF-спецификаций используем синтаксис Turtle:

```
prefix vCard:
  <http://www.w3.org/2001/vcard-rdf/3.0#> .
prefix rdf:
  <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
prefix : <#> .
<http://somewhere/MattJones/>
  vCard:FN "Matt Jones" ;
  vCard:N [ vCard:Family "Jones";
           vCard:Given "Matthew" ] .
```

Пространства имён, определяемые в RDF как словари, отображаются в миры фреймов. Мир фреймов в языке СИНТЕЗ представляется с помощью фрейма с указанием принадлежности его метаклассу world, определяющему семантику этого фрейма как мира. Внутри мира располагаются спецификации входящих в него фреймов.

```
{ vCard;
  in: world;
  ...
}
```

Значение слота может включать вложенный фрейм в соответствии с возможностью задавать в RDF пустые узлы.

С определённым субъектом триплетов связывается фрейм с соответствующим именем. Набор триплетов с общим субъектом отображается в набор слотов одного фрейма. Предикатам, связанным с этим субъектом, соответствуют слоты фрейма. Объектам – значения слотов. Если тип

значения не определён, он назначается как тип Frame, так как фреймы формируют произвольные значения. Набор триплетов, различающихся объектами, отображается в слот фрейма с несколькими значениями.

```
{ MattJones;
  in: frame;
  FN: Matt Jones;
  N: { in: frame; Family: Jones; }
    { in: frame; Given: Matthew; }
}
```

В графе имеются неименованные узлы, являющиеся объектами предиката N, которые выражаются встроенными фреймами без идентификатора.

Таким образом, данные в базовой модели RDF могут представляться как база фреймов. В RDF словари расширяют модель и могут приносить определённую семантику элементам RDF-модели. Поэтому определение схем над RDF-графами, исполняемое в модели RDF Schema, несёт в себе семантику определения структуры и влияет на отображение RDF Schema в язык СИНТЕЗ.

#### 4.2 Отображение модели RDF-Schema в объектную модель

Описания в модели RDF Schema отображаются в объектную модель языка СИНТЕЗ.

Сложности отображения RDF в объектную модель (как и реляционную) связаны, в основном, с конфликтом парадигм открытого мира в случае RDF и закрытого мира в случае объектной модели [9]. Во-первых, описываемые ресурсы могут принадлежать нескольким RDF-классам одновременно. Большинство объектных моделей исключает такую возможность, определяя принадлежность объекта строго одному типу. Во-вторых, свойства в RDF являются самостоятельными сущностями, которые могут использоваться в разных классах и ресурсах. Их связь с классами определяется только указанием классов домена и области значений свойств. Тем временем, атрибуты типов в объектной модели обычно жёстко связаны с типами, а также наследуются от супертипов. В-третьих, структура экземпляров в RDF не ограничивается спецификациями классов. В большинстве объектных моделей объект жёстко соответствует структуре, описанной в типе, которому он принадлежит. Средства языка СИНТЕЗ позволяют избежать указанных проблем, благодаря разделению в языке на классы как множества и абстрактные типы данных как интенциональные описания, а также возможности использовать типовые выражения для их комбинации.

Определения RDF Schema определяют схему RDF-данных посредством задания ресурсов, классов, свойств и их ограничений.

```

prefix rdf:
  <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix rdfs:
  <http://www.w3.org/2000/01/rdf-schema#>
prefix vcard:
  <http://www.w3.org/2001/vcard-rdf/3.0#>
vCard:Individual rdf:type rdf:Class .
vCard:FN rdf:type rdf:Property;
  rdfs:domain vCard:Individual .
vCard:NPROPERTIES rdf:type rdfs:Class .
vCard:N rdf:type rdf:Property;
  rdfs:domain vCard:Individual;
  rdfs:range vCard:NPROPERTIES .
vCard:Family rdf:type rdf:Class;
  rdfs:subClassOf vCard: NPROPERTIES .
vCard:Given rdf:type rdf:Class;
  rdfs:subClassOf vCard: NPROPERTIES .

```

Ресурс RDF Schema отображается во фрейм языка СИНТЕЗ. Класс RDF Schema отображается в класс языка СИНТЕЗ с типом экземпляров.

```

{ individual;
  in: class;
  instance_section:
  {
    FN: {set_of: Frame};
    N: {set_of: nproperties};
  };
},
{ nproperties;
  in: class;
  instance_section: Frame
},
{ family;
  in: class;
  superclass: nproperties
  instance_section: Frame
},
{ given;
  in: class;
  superclass: nproperties
  instance_section: Frame
};

```

Свойства RDF Schema отображаются в общем случае в метакласс ассоциаций, тем самым создавая соответствующую самостоятельную сущность.

```

{ n;
  in: association, metaclass;
  instance_section:
  { association_type: {0, inf},{0, inf};
    domain: individual;
    range: nproperties;
  }
}

```

В этом случае атрибут типа определяется с указанием метакласса ассоциаций:

```

N: {set_of: NPROPERTIES};
  metaslot
    in: n;
  end

```

Свойства domain и range соответствуют определению домена и диапазона значений в метаклассе ассоциаций. Если домен определяется как единственный класс, вместо метакласса ассоциаций возможно определять обычный атрибут в типе.

Свойства subclass и subproperty соответствуют отношению подкласса между классами канонической модели и метаклассами ассоциаций соответственно.

Конструкции Bag, Seq, List отображаются в коллекции канонической модели.

В случае присутствия в ресурсе свойства type по отношению к классу RDF Schema ресурс отображается в экземпляр класса, и фрейм, представляющий его, рассматривается как объект типа экземпляров данного класса.

```

{ MattJones;
  in: individual;
  FN: Matt Jones;
  N: { Jones; in: family;}, { Matthew; in: given;};
}

```

В языке RDF Schema определяется набор правил вывода, который должен использоваться над спецификациями. Некоторые из этих правил представимы в виде правил и инвариантов на языке формул канонической модели, а некоторые – предполагаются языком СИНТЕЗ и не требуют дополнительной спецификации. Однако разные реализации RDF ограничиваются использованием подмножества правил, указанных в стандарте RDF. В зависимости от набора реализованных правил вывода в информационном ресурсе, они должны

отображаться или не отображаться в спецификацию на языке СИНТЕЗ. Также необходимость их спецификации зависит от уровня их применения: они могут применяться при обновлении данных в информационном ресурсе, чего посредник СИНТЕЗ не предполагает делать; либо их применение зашито в обработчике запросов ресурса, тогда в спецификации достаточно инвариантов, проверяющих целостность данных.

### 4.3 Отображение запросов в язык SPARQL

В модели RDF нет собственного языка манипулирования данными. Для запросов над базами RDF-документов обычно используется язык SPARQL. Запросы на языке SPARQL представляют собой конъюнкции триплетов, на месте субъектов, предикатов и объектов в которых могут быть определённые значения или переменные. Пример запроса:

```
prefix vcard: <http://www.w3.org/2001/vcard-rdf/3.0#>
SELECT ?x ?fname ?gname
{
  ?x vcard:N ?vc .
  ?vc vcard:Family ?fname .
  ?vc vcard:Given ?gname .
}
```

Для языка манипулирования данными строится обратное отображение: из языка СИНТЕЗ в запрос на языке SPARQL – для возможности переписывания запросов к информационным ресурсам. В соответствии с возможностями языка запросов SPARQL, в них возможно отображение формул языка СИНТЕЗ в конъюнктивной форме над предикатами классов и условиями. Для формирования приведённого выше запроса на языке SPARQL в посреднике должен быть задан запрос на языке формул:

```
q([x, fname, gname]) :-
  individual(x),
  ex y (family(y) & is_in(y, x.N) & fname=y),
  ex z (given(z) & is_in(z, x.N) & gname=z)
```

В случае отображения RDF во фреймы канонической модели в качестве запросов применяются формулы канонической модели над базой фреймов [8]. В них фигурируют операции над фреймами и не используется типизация.

Таким образом, возможно отображение данных в модели RDF в базу фреймов языка СИНТЕЗ. При наличии спецификации схемы данных RDF она отображается в объектную спецификацию. Формулы ограниченного вида в запросах и программах посредника преобразуются в запросы на SPARQL.

### 4.4 Учёт производных моделей данных и логического вывода

Востребованность модели данных RDF возростала последовательно в течение ряда лет. Устойчивым интересом к себе она обязана тем, что стала базовой моделью для семантического веба. Тем временем, хранилища RDF долгое время были слабо масштабируемы до развития принципов вертикального хранения, что на интенсивности использования модели сказывалось негативно. Однако производная модель данных OWL, благодаря своей выразительности и возможности автоматического вывода в задаче включения классов, стала использоваться для описания схем баз данных и знаний.

Экземпляры классов в модели OWL представимы в модели RDF. Поэтому в случае, когда RDF-данные ограничены схемой, описанной на языке OWL, это существенно влияет на семантическое отображение таких RDF-данных в другие модели, в частности, в язык СИНТЕЗ.

Отображение моделей данных OWL и OWL 2 в модель, определяемую языком СИНТЕЗ, было подробно рассмотрено ранее [18].

Логический вывод, предполагаемый в модели OWL, учитывается при обновлении данных. В этом случае, правила вывода далее не должны учитываться в отображении моделей.

## 5 Заключение

Исследованы модели данных RDF и RDF Schema, определены принципы отображения RDF-спецификаций во фреймовую модель унифицирующей модели предметных посредников, определены принципы отображения спецификаций RDF Schema в объектную каноническую модель, описаны возможности запросов в канонической модели, отображаемых в язык запросов SPARQL, используемый для RDF-спецификаций и производных от них моделей. Результаты работы могут быть использованы для семантического отображения RDF-описаний в предметную область посредников и доступа к RDF-данным в них.

## Литература

- [1] OWL Web Ontology Language Reference. M. Dean, G. Schreiber (Eds.), W3C Recommendation. – W3C, 2004. – URL: <http://www.w3.org/TR/2004/REC-owl-ref-20040210/>.
- [2] RDF vocabulary description language 1.0: RDF schema. D. Brickley, R.V. Guha (Eds.), W3C Recommendation. – W3C, 2004. – URL: <http://www.w3.org/TR/2004/REC-rdf-schema-20040210/>.
- [3] SPARQL Query Language for RDF. E. Prud'hommeaux, A. Seaborne (eds.), W3C Recommendation. – W3C, 2008. – URL: <http://www.w3.org/TR/rdf-sparql-query/>

- [4] D. Beckett, J. Grant. SWAD-Europe Deliverable 10.2: Mapping Semantic Web data with RDBMSes. – 2003. – URL: [http://www.w3.org/2001/sw/Europe/reports/scalable\\_rdbms\\_mapping\\_report](http://www.w3.org/2001/sw/Europe/reports/scalable_rdbms_mapping_report).
- [5] Chebotko A., Lu S., Fotouhi F. Semantics preserving SPARQL-to-SQL translation // *Data & Knowledge Engineering*. – 2009. – Т. 68. – №. 10. – С. 973-1000.
- [6] Erling O. Virtuoso, a Hybrid RDBMS/Graph Column Store // *Data Engineering*. – 2012. – С. 3.
- [7] Kalinichenko L.A., Stupnikov S.A. Heterogeneous information model unification as a pre-requisite to resource schema mapping. // A. D’Atri and D. Sacca (eds.), *Information Systems: People, Organizations, Institutions, and Technologies* (Proc. of the V Conference of the Italian Chapter of Association for Information Systems itAIS). – Berlin-Heidelberg: Springer Physica Verlag, 2009. – P. 373-380.
- [8] Kalinichenko L.A., Stupnikov S.A., Martynov D.O. SYNTHESIS: a Language for Canonical Information Modeling and Mediator Definition for Problem Solving in Heterogeneous Information Resource Environments. – Moscow: IPI RAN, 2007. – 171 p.
- [9] Oren E. et al. ActiveRDF: Object-oriented semantic web programming. // *Proceedings of the 16th international conference on World Wide Web*. – ACM, 2007. – С. 817-824.
- [10] Sahoo S. S. et al. A survey of current approaches for mapping of relational databases to rdf // *W3C RDB2RDF Incubator Group Report*. – 2009.
- [11] K. Wilkinson, C. Sayers, H. A. Kuno, D. Reynolds. Efficient RDF Storage and Retrieval in Jena2. // *In Semantic Web and Databases Workshop*. P. 131–150, 2003.
- [12] Yu L. *Linked open data // A Developer’s Guide to the Semantic Web*. – Springer Berlin Heidelberg, 2011. – С. 409-466.
- [13] M. A. Bornea, et al. Building an Efficient RDF Store Over a Relational Database. // *Proc. of the 2013 ACM SIGMOD International Conference on Management of Data*. – New York, 2013. – P. 121-132.
- [14] Брюхов Д.О., Вовченко А. Е., Захаров В.Н., Желенкова О.П., Калиниченко Л.А., Мартынов Д.О., Скворцов Н.А., Ступников С.А. Архитектура промежуточного слоя предметных посредников для решения задач над множеством интегрируемых неоднородных распределенных информационных ресурсов в гибридной грид-инфраструктуре виртуальных обсерваторий // *Информатика и ее применения*. – М., 2008. – Т. 2, Вып. 1. – С. 2-34.
- [15] Н. А. Скворцов. Отображение моделей данных NoSQL в объектные спецификации // *RCDL’2012. – CEUR Workshop Proceedings, 2012. – Т. 934. – С. 53-62*.
- [16] С. А. Ступников. Унификация модели данных, основанной на многомерных массивах, при интеграции неоднородных информационных ресурсов // *RCDL’2012. – CEUR Workshop Proceedings, 2012. – Т. 934. – С. 42-52*.
- [17] С. А. Ступников. Отображение графовых моделей данных в каноническую информационную модель при интеграции неоднородных информационных ресурсов // *RCDL’2013. – Готовится к печати*.
- [18] С. А. Ступников, Н. А. Скворцов. Взаимное отображение канонической информационной модели и языка OWL 2 // *RCDL’2010. – Казань: КФУ, 2010. – P. 392-398*.

## Mapping of RDF Data Model into the Canonical Model of Subject Mediators

Nikolay A. Skvortsov

The RDF data model becomes more and more widely used to identify, describe and interrelate resources of any kind in the open information environment. During the problem solving, there exists a need to involve data represented in this model in the process of multiple heterogeneous information resource integration. The paper analyzes approaches for mapping of the RDF data model expressed by its accompanying languages (RDF Schema and SPARQL) into the SYNTHESIS language. The SYNTHESIS language is used as a unifying canonical model for integration of multiple heterogeneous information resources in information systems created using subject mediation technique. An approach for RDF data model unification is proposed that takes into account discovery of RDF data semantics and schema.