

# Формирование глоссария по онтологическому моделированию: состояние и перспективы развития<sup>1</sup>

Н. А. Скворцов

Институт проблем информатики РАН  
Россия, г. Москва, 117333, ул. Вавилова, 44/2  
naskv@ipi.ac.ru

**Аннотация.** Глоссарий по онтологическому моделированию разрабатывается как русскоязычный словарь терминов с определениями и семантическими отношениями. В статье описаны цели и требования, предъявляемые к разрабатываемому глоссарию, структура его статей, охват терминами предметных областей, смежных с областью онтологического моделирования, принципы коллективной разработки статей глоссария. Представлено состояние его разработки на момент написания статьи, описаны частные вопросы терминологии, решаемые с помощью глоссария, описаны перспективы его развития.

**Ключевые слова:** онтологическое моделирование, глоссарий, терминология.

## 1 Введение

Одним из решений, принятых участниками первого симпозиума «Онтологическое моделирование» в 2008 году [1], было намерение создать русскоязычный словник терминов предметной области онтологического моделирования. Этому решению послужили выявленные в ходе симпозиума различия в понимании предметной области, в толковании основополагающих понятий, а также в переводе на русский язык одних и тех же англоязычных терминов.

С помощью словника предполагалось достигнуть следующие цели:

---

<sup>1</sup> Работа выполнена при финансовой поддержке РФФИ (проекты 08-07-00157-а, 10-07-00342-а, 11-07-00402-а) и Программы фундаментальных исследований Президиума РАН № 15, проект 4.2.

- согласование понимания терминов внутри сообщества симпозиума «Онтологическое моделирование»;
- формирование русскоязычного варианта терминологии в области онтологического моделирования;
- выработка рекомендаций по использованию терминов специалистами в онтологическом моделировании.

В. Ш. Рубашкиным была составлена начальная заготовка словника и обозначено, что желаемым результатом этой работы мог бы стать двуязычный словарь терминов, снабженных краткими определениями. Таким образом, начальный вариант словника послужил отправной точкой разработки глоссария.

Глоссарий по онтологическому моделированию создаётся как двуязычный (русскоязычный с английскими аналогами терминов) систематизированный словарь терминов предметной области, снабжённых определениями (вариантами определений) и существенными семантическими отношениями. Его разработка должна проводиться создаваемой для этого рабочей группой посредством интерактивного интернет-сайта.

В разделе 2 описаны терминологические проблемы в области онтологического моделирования и связанные с ними требования к создаваемому глоссарию. Раздел 3 описывает охватываемые области, смежные с онтологическим моделированием, которые должны быть представлены терминами в глоссарии. В разделе 4 приводится структура статей глоссария и принципы их заполнения. В разделе 5 приведены сведения о реализации поддержки разработки глоссария на сайте. Раздел 6 подробно описывает этапы построения глоссария рабочей группой. Раздел 7 посвящён состоянию глоссария на время написания статьи и некоторым вопросам терминологии, которые поднимаются в глоссарии. В разделе 8 перечислены обстоятельства, вызывающие трудности при разработке глоссария, и расставлены акценты его анализа и развития.

## **2 Проблемы терминологии онтологического моделирования**

Предметная область онтологического моделирования развивается на пересечении ряда областей знания, имеющих собственную богатую историю развития и научных исследований. Онтологическое моделирование базируется на знаниях из областей философии, лингвистики, логики, концептуального моделирования. Терминология этих областей используется в онтологическом моделировании. Однако существует также и тенденция применять терминологию, специфическую для онтологического моделирования, к различным областям знаний. В частности, в концептуальном моделировании баз данных и проектировании информационных систем изменение терминологии на

онтологическую во многом продиктовано удобством применения языка представления онтологий OWL (Web Ontology Language) [2] в качестве языка определения данных. В области представления знаний использование терминологии онтологического моделирования объясняется применением того же языка, так как он основан на дескриптивных логиках и удобен в задаче классификации фактов. В лингвистике онтология фигурирует, в первую очередь, как словарь лингвистических термов. В результате, термины, исторически присущие перечисленным областям, нередко подменяются терминами из области онтологического моделирования. И наоборот, с терминами области онтологического моделирования связывается семантика, присущая этим областям. При этом увеличиваются разночтения в семантике терминов онтологического моделирования в разных сообществах. На терминологию в целом в области онтологического моделирования также оказывают большое влияние частные терминологические решения, принятые в конкретных средствах моделирования, в том числе, фактически не связанных с онтологическим моделированием.

В целом, терминологию в области онтологического моделирования до сих пор трудно назвать устоявшейся. В частности, русскоязычная терминология, помимо всего прочего, зависит от англоязычной как международной, и разночтения в ней возникают в результате разных переводов терминов с английского языка на русский.

### **3 Охват терминологии глоссария**

Для того чтобы разобраться в обозначенных выше проблемах терминологии в области онтологического моделирования, в разрабатываемый глоссарий по онтологическому моделированию необходимо включать термины не только из области собственно онтологического моделирования, но и из областей знания, составляющих её инфраструктуру, давая терминам определения в контексте этих областей.

Для этого в глоссарии введены рубрики, соответствующие областям знания или некоторым общим признакам терминов, принадлежащих этим рубрикам. По состоянию на время написания данной статьи в глоссарии присутствуют следующие рубрики:

- «Алгоритмы» – рубрика, содержащая термины, связанные со сложностью алгоритмов;
- «Деятельность» – рубрика, содержащая термины, относящиеся к разновидностям действий агентов или неразрывно связанные с действиями агентов;

- «Информационные системы» – рубрика, содержащая термины, относящиеся к разновидностям информационных систем или специфические для области проектирования информационных систем;
- «Информационный поиск» – рубрика, содержащая термины, специфические для систем информационного поиска;
- «Концептуальное моделирование» – подрубрика рубрики «Моделирование», содержащая термины, связанные с концептуальным моделированием баз данных и информационных систем;
- «Лингвистика» – рубрика, содержащая термины, пришедшие из области лингвистики;
- «Логика» – широкая рубрика, содержащая термины, относящиеся к основам математической логики, задачам логического вывода;
- «Моделирование» – рубрика, содержащая термины, общие для различных видов моделирования;
- «Онтологическое моделирование» – подрубрика рубрики «Моделирование», содержащая термины, непосредственно принадлежащие области онтологического моделирования и специфические только для неё;
- «Организация информации» – рубрика, содержащая термины, связанные с различными видами представления информации, способами её упорядочения;
- «Проектирование онтологий» – подрубрика рубрики «Деятельность», содержащая термины, относящиеся к деятельности, производимым агентами в процессе проектирования онтологий;
- «Семантика» – рубрика, содержащая термины, относящиеся к области семантики термов в естественных и искусственных языках;
- «Согласование онтологий» – подрубрика рубрики «Деятельность», содержащая термины, относящиеся к деятельности, производимым над несколькими онтологиями с целью из согласования.

В процессе развития глоссария состав рубрик может быть изменён и дополнен.

С помощью введения рубрик достигается несколько целей. Во-первых, разграничивается терминология онтологического моделирования и терминология, исторически присущая другим областям знания. Во-вторых, выясняется преемственность терминов из разных областей знания в области онтологического моделирования.

Каждый термин должен принадлежать одной или нескольким рубрикам. При этом определение термину даётся в контексте рубрики, из которой он исторически пришёл. Если в контекстах разных рубрик (в разных областях знания) семантика терминов отличается либо уточняется в производной рубрике, то вводятся разные термины как омонимы с самостоятельными определениями либо сами имена терминов изменяются в соответствии с их особенностями в разных контекстах.

#### 4 Описание терминов глоссария

Всякий термин в глоссарии должен иметь вербальное определение. При определении терминов глоссария необходимо руководствоваться принципами, которые позволяют контролировать их качество).

- Вербальные определения должны действительно определять, что означают термины. Определения должны описывать существенные свойства терминов, т. е. подчёркивать существенные особенности смысла термина. Если для определения термина существенно указать его экземпляры или разновидности, то они могут быть перечислены.
- Предпочтительны достаточно краткие и точные определения.
- Необходимо избегать цикличности определений, т. е. взаимного определения терминов друг через друга.
- Необходимо избегать двусмысленности в текстах определений.
- В глоссарии должны быть отдельно определены все значимые в данной области термины, встречающиеся в определениях других терминов.
- Приоритет отдаётся определениям, взятым из представительных источников. Однако определения могут быть изменены для более точного определения терминов либо составлены независимо от принятых в литературе определений.
- В случае цитирования определений необходимо указывать источник определения.

Помимо вербального определения, статьи, описывающие термины глоссария, могут описывать связи между терминами и включать разъяснения, касающиеся терминов. Структура полей в статьях глоссария соответствует шаблону, приведённому ниже.

##### 1. Вербальное определение термина:

- имя термина на русском языке;
- вербальное определение термина на русском языке либо нумерованный список вариантов определения одного и того же термина (но не омонимов этого термина);
- англоязычный аналог термина;
- англоязычное вербальное определение термина, если определение взято из англоязычного источника;
- источники (публикация или URL) определения термина.

##### 2. Связанные термины (определение существенных семантических отношений данного термина с другими терминами глоссария):

- список синонимов (они не определяются отдельными статьями);
- более общий термин (родовой);
- указатель на список омонимов термина в разных областях знания;
- части и целое;
- список рубрик, которым принадлежит термин;
- другие семантические отношения.

### 3. Дополнительные описания:

- разъяснения и рекомендации об интерпретации термина и употреблении его в предметной области;
- комментарии к статье, которые необходимы в процессе разработки глоссария.

## 5 Автоматизация работы над глоссарием

Для разработки и представления глоссария в электронном виде создан wiki-ресурс [3], в котором каждый термин глоссария определён на отдельной странице (рис. 1). Структура страниц соответствует описанному в предыдущем разделе шаблону статей глоссария. Вхождения других терминов глоссария в определение текущего термина отмечаются ссылками на страницы с определениями этих терминов. Семантические отношения с другими терминами обозначаются видом отношения и ссылками на страницы терминов, состоящих в данном отношении с текущим термином. В случае если термин имеет омонимы, указывается ссылка на страницу многозначного термина. Страница многозначного термина содержит перечень всех омонимов с данным именем и ссылки на их страницы, сами омонимы описываются на разных страницах. Рубрики, которым принадлежат термины, реализуются с помощью wiki-категорий.

Для совместной работы формируется рабочая группа по созданию глоссария. Участники рабочей группы соглашаются об общих принципах работы над глоссарием. Работа по созданию глоссария проводится на основе wiki-ресурса. При разработке принципов функционирования рабочей группы был проанализирован опыт коллективной разработки глоссария в проекте NeOn [4] и взяты на вооружение некоторые средства, используемые при этом.

Для поддержки обсуждения термина рабочей группой установлена система Cicero [5]. Она позволяет формулировать вопросы (проблемы), связанные со страницами терминов, предлагать решения, оценивать и аргументировать предложенные решения, принимать решения на основе обсуждения, устраивать голосование. Ссылка на проект разработки глоссария в системе Cicero располагается в разделе «Проекты и задачи» wiki-ресурса. Также для обсуждения вопросов, требующих пространных и неформальных дискуссий, установлен форум.

Добавить вопрос к данной странице описания термина, чтобы инициировать по нему дискуссию, а также просмотреть список связанных с данной страницей вопросов можно в разделе «Обсуждение» данной страницы (см. рис. 1).

статья	обсуждение	просмотр	история
--------	------------	----------	---------

---

## Онтология верхнего уровня

---

### Определение

Онтология верхнего уровня - **онтология**, содержащая **понятия** которые являются метапонятиями, порождающими, абстрактными, философскими **понятиями**, достаточно общими для того, чтобы их можно было употреблять в широком спектре **предметных областей**.

**Англоязычный аналог:** Top-Level Ontology (Upper Ontology, Foundation Ontology) is limited to concepts that are meta, generic, abstract and philosophical, and therefore are general enough to address (at a high level) a broad range of domain areas.

---

### Источники

The IEEE Standard Upper Ontology. <http://suo.ieee.org/>

---

### Связанные термины

Более общее понятие: [Онтология](#)

---

### Дополнительные описания

---

### Комментарии

---

Категории: [Список терминов](#) | [Онтологическое моделирование](#)

Рис. 1 Статья глоссария в wiki-ресурсе

## 6 Этапы работы над глоссарием

Процесс создания глоссария включает следующие этапы, регламентирующие последовательность действий участников рабочей группы по его разработке.

### 6.1 Этап 1 – Создание первичного варианта глоссария

Первичный вариант глоссария создаётся силами нескольких участников рабочей группы до начала коллективной разработки. На данном этапе определяются охватываемые области знания. Собираются существенные в данных областях термины. Эти термины определяются вербально и включаются в глоссарий. После этого рекурсивно до некоторого насыщения в глоссарий включаются существенные термины, фигурирующие в определениях.

Полученный вариант глоссария корректируется несколькими экспертами рабочей группы и выставляется в wiki-ресурсе в виде множества страниц, описывающих отдельные термины.

Данный этап создания глоссария на время написания статьи был выполнен. С использованием заготовки словника, составленной В. Ш. Рубашкиным, автором данной статьи был собран первичный вариант глоссария (около 200 терминов с определениями). Данный глоссарий был скорректирован Л. А. Калиниченко и М. Р. Коголовским. Автор выражает благодарность им за подробные комментарии и правки определений. Результат работы на данном этапе выложен на wiki-ресурсе [3].

В конце этапа определяется состав рабочей группы по созданию глоссария. Все участники оповещаются о начале коллективной работы над глоссарием, принципах и этапах работы.

## **6.2 Этап 2 – Комментирование и пополнение глоссария**

Участники рабочей группы регистрируются на сайте, знакомятся с первичным вариантом глоссария на wiki-ресурсе, дают комментарии к статьям и дополняют глоссарий недостающей информацией.

Важно, чтобы при знакомстве с первичным вариантом глоссария участники уделили внимание каждому описанному термину и прокомментировали максимально большое количество статей глоссария при возникновении любых вопросов, оценивая при этом:

- уместность включения терминов в глоссарий;
- точность имён и определений терминов;
- приемлемость перевода терминов и определений с английского языка;
- корректность отношений с другими терминами (синонимы, более общие термины, омонимы, рубрикация и так далее).

Также каждый участник оценивает полноту глоссария и дополняет глоссарий недостающей информацией:

- создаёт новые страницы с терминами, которые, по его мнению, должны быть в глоссарии, но отсутствуют, используя при этом общий шаблон построения статей глоссария, описанный в разделе 4, и указывая русскоязычные и англоязычные варианты терминов, определение, источник определения и, по возможности, отношения с другими терминами;
- составляет комментарии к существующим страницам терминов, с альтернативными определениями и указанием их источников, если, по мнению участника, существуют другие значения или более уместные определения термина;
- указывает в комментариях отсутствующие связи терминов, принадлежность к рубрикам.



При комментировании или дополнении глоссария участники рабочей группы должны поставить свою подпись (имя пользователя на сайте) для организации дальнейшего обсуждения глоссария. Ответственным лицом производятся правки для однородного представления исправленных и новых статей глоссария.

### **6.3 Этап 3 – Обсуждение вопросов и принятие решений**

Этап проводится в автоматизированной системе обсуждения Cicero [5], функционирующей в wiki-ресурсе. Ответственное лицо оповещает участников рабочей группы о переходе к данному этапу после окончания комментирования и пополнения глоссария отдельно каждым участником.

До начала этапа по комментариям участников, составленным на предыдущем этапе, ответственным лицом в системе регистрируется начальный список вопросов для обсуждения. В системе Cicero каждый участник рабочей группы может:

- регистрировать свои вопросы, относящиеся к статьям глоссария (это удобно делать в разделах «Обсуждение» страниц wiki-ресурса);
- предлагать свои варианты решения зарегистрированных в системе Cicero вопросов;
- оценивать варианты решения, подтверждая их или опровергая, с описанием примеров, оценок по какому-либо критерию, словесных обоснований.

Активность ожидается от каждого участника в составлении вопросов, предложении вариантов решения каждого вопроса и оценке каждого предложенного варианта решения.

Решение по результатам обсуждения вопросов в зависимости от ситуации выносится одним из следующих способов:

- ответственным лицом на основе предложенных решений (при очевидном решении);
- по преобладающим оценкам предложенных решений – отрицательным и положительным (в редких случаях, так как обычно участники делают только отрицательные оценки, если не согласны, а положительных не делают);
- проведением голосования по выбору предложенных решений (если обсуждения продолжаются без видимого выхода или завершаются без возможности достижения консенсуса).

Манипуляции над статьями глоссария в целом в процессе их обсуждения производятся по следующим принципам:

- если комментарии к статье глоссария отсутствуют или в целом положительные, нет открытых вопросов по данному термину, то определение рассматривается как окончательное;

- если комментарии отсутствуют или в целом положительные, но были вопросы по определению, то до тех пор, пока вопросы не перестанут появляться, предлагаются варианты изменений в определении (-ях);
- если комментарии в целом отрицательные, ставится вопрос о переопределении термина, где в качестве решений могут предлагаться: другие определения, удаление термина, введение в глоссарий нескольких альтернативных определений одного термина, введение омонимов термина, выбор другой гранулярности терминов в предметной области и другое;
- при необходимости неформальных обсуждений создаются дискуссии в форуме, дискуссии продолжаются, пока результаты не будут достигнуты;
- сложные случаи откладываются для обсуждения при встрече на предстоящем симпозиуме.

## 7 Состояние разработки глоссария

В разделе 3 перечислены рубрики, представляющие собой области знания, термины которых должны быть включены в глоссарий по онтологическому моделированию. Но набор терминов не определяется тем, какие рубрики присутствуют в глоссарии. Напротив, перечисленные рубрики и состав вошедших в них терминов – результат того, какие вопросы терминологии, используемой в онтологическом моделировании, решались при составлении глоссария. Ниже перечислены некоторые вопросы, которые отражены терминами в глоссарии в его состоянии на время написания статьи.

### 7.1 Специфические термины онтологического моделирования

*Онтологическое моделирование* – это процесс построения, развития, обработки и использования онтологии предметной области. В глоссарии предлагается зафиксировать несколько альтернативных определений *онтологии* как явной спецификации концептуализации предметной области, как совокупности понятий предметной области и отношений между ними [6] и как словаря предметной области с логической теорией над ним [7]. Эти определения рассматривают термин *онтология* с разных ракурсов и дополняют друг друга, другие определения являются, в основном, вариациями приведённых определений.

*Понятие* – мысленное представление, которое рассматривается как эквивалент сущностей для некоторых целей [8]. Под термином *концептуализация* понимается абстрактное осмысление предметной области, при котором она разделяется на понятия [9]. С

концептуализацией предметной области связан также термин *гранулярность понятий*.

Определение термина «онтология» связано с термином *онтологическое обязательство* (данный перевод термина "ontological commitment" предпочтителен, так как он принят в философском сообществе), означающим либо обязательство логической теории к существованию в предметной области сущностей определённого вида [10], либо соглашение агентов использовать словарь непротиворечивым образом по отношению к теории, специфицируемой онтологией [11].

В глоссарии описаны термины, связанные с особыми видами онтологий: *формальная онтология*, *онтология верхнего уровня*, *метаонтология* [12]. Отличие двух последних терминов состоит в том, что метаонтологии специфицируют множество онтологий, определяя общие свойства их понятий, а онтологии верхнего уровня содержат понятия, достаточно общие для того, чтобы их можно было употреблять в широком спектре предметных областей.

Определены также термины, связанные с технической инфраструктурой онтологического моделирования: *библиотека онтологий*, *редактор онтологий*, *семантический веб*.

## 7.2 Термины, используемые в концептуальном моделировании

*Концептуальное моделирование* занимается описанием сущностей, их структур, характерных соотношений между ними и их поведения в предметной области [13]. Эта область имеет большую историю развития, связанную, в первую очередь, с базами данных.

Важно определить термины, связанные с языками, на которых происходит описание сущностей. Заслуживающим внимания является термин *модель данных*. Исторически в области концептуального моделирования под моделью данных понимается не результат моделирования части реального мира, а средства и правила моделирования. Конкретными моделями данных являются *реляционная модель данных*, *иерархическая модель данных* и другие. Модели данных определяются языками моделирования, а именно *языком определения данных* и *языком манипулирования данными*. Результату же концептуального моделирования, т. е., описанию предметной области на языке концептуального моделирования, соответствует термин *концептуальная схема*. Вне зависимости от того, в какой области используется термин «модель данных», его необходимо использовать в изначальном исторически принятом смысле.

Таким же образом предпочтительно использовать и термин *информационная модель*, являющийся более общим по отношению к термину «модель данных» и означающий язык представления

информационных сущностей, связей, правил и операций для спецификации семантики информации. К сожалению, в различных источниках последних лет чаще можно встретить информационную модель в значении результата моделирования части реального мира с помощью информации.

Для того чтобы избежать путаницы в производных терминах, необходимо определить, что *концептуальная модель данных* – это модель данных (т. е. язык концептуального моделирования), используемая в качестве инструмента для представления концептуальной схемы базы данных. А *концептуальная модель предметной области* – это её абстрактное описание, независимое от аспектов реализации систем, определяющее концептуальную структуру и поведение сущностей в этой предметной области, формально выраженное в виде концептуальной схемы [13].

Процессы, связанные с преобразованием информационных моделей, отражаются в глоссарии терминами *отображение моделей* как построение правил преобразования исходной модели в целевую и *унификация моделей* как приведение множества информационных моделей к одной, которую при этом называют *канонической*.

Термин *модель*, общий для разных направлений моделирования, соответствует упрощённому абстрактному представлению реальности. *Схема* представляет модель на определённом языке моделирования. Определены термины *язык моделирования*, *выразительная мощьность языка*, *спецификация*.

Термин *сущность* соответствует самостоятельно существующим в предметной области единицам. Термин *объект* описывает пассивные единицы, которыми можно оперировать, имеющие состояние, поведение, идентичность. Конкретное состояние объекта описывается *значением*. Область допустимых значений и операции над объектами описываются *типами*. Термин *абстрактный тип данных* подразумевает, что тип является структурой, включающей спецификацию поведения и состояния объектов. Множества сущностей с каким-либо одинаковым свойством или одинакового типа определяются *классами*. Понятие типа отличается от понятия класса тем, что концентрируется на описании множества общих свойств сущностей (*интенционала*), а не на определении множества сущностей (*экстенционала*). С типами связаны термины *подтип* и *супертип*, с классами – *подкласс* и *суперкласс*, определяемые отношениями обобщения/специализации между сущностями. Термин *категория* используется для фундаментальных классов, определяющих разновидности сущностей. В частности, Аристотель выделил 10 категорий, в которые попадают любые сущности.

Среди терминов, отвечающих за организацию информации, также определены термины *система классификации* как набор классов для отнесения по ним сущностей в соответствии с некоторыми критериями,

*таксономия* как любая упорядоченная система классификации, *иерархия* как множество сущностей, частично упорядоченное по какому-либо отношению. *Кластер* – объединение нескольких однородных элементов, которое может рассматриваться как самостоятельная единица, обладающая определёнными свойствами. Сущности могут быть описаны посредством фреймов. Термин *фрейм* определяется как структура данных для описания шаблонных ситуаций. Элементами фреймов являются *слоты*, которые, в свою очередь, имеют значения.

Термин *метаданные* означает данные о данных, т. е. описание формы и содержания документов, объектов, сервисов [14]. С метаданными, сопровождающими данные, обычно связывают термин *семантическая аннотация*.

Необходимо провести различие между терминами «значение», «экземпляр», «индивид». *Значением* является состояние, которое принимает объект определенного типа или переменная. Об *экземпляре* говорят, когда существует класс, представителем которого является сущность. Термин *индивид* используется для конкретных сущностей с собственной идентификацией. В отличие от объекта, который рассматривается агентами и которым оперируют, индивид существует независимо от его рассмотрения.

В глоссарии важно разобраться в употреблении терминов для разного рода связей. *Связь* (relationship) – это любой способ соотнесения двух или более сущностей друг с другом. Термин *отношение* (relation) следует использовать для конкретных связей конкретного вида, в математическом смысле отношению соответствует бинарный предикат или предикат более высокой арности, в этом же смысле термин «отношение» используется в базах данных. *Ассоциацией* называют отношение, означающее некоторую смысловую связь между классами. *Фактом* является утверждение об индивиде или сущности, в частности, связь, имеющая место в конкретном случае.

Часто для обозначения конкретных связей используют термин *свойство*. Однако в отличие от отношения *свойством* (property) называется одноместный предикат [15], другими словами, свойства не описывают связи. *Свойство* – черта, которой обладает сущность, даже если мы не осведомлены об этом факте. *Атрибут* – это свойство, характерная черта, которую мы присваиваем сущности. *Роль* – свойство, характеризующее, каким образом нечто участвует в зависящих от обстоятельств событиях или в существующем положении дел. Роль как свойство «являться чем-либо» бывает связана с отношением «X имеет Y в качестве R», так как обратное отношение – «X является R по отношению к Y», где R – свойство, являющееся ролью.

Со спецификациями поведения сущностей в концептуальном моделировании связаны следующие термины. *Операции* означают действия, выполняемые над одним или более объектами с целью

получения нового результата. Определены термины *предусловие* и *постусловие*, вместе формализующие спецификацию операции. У термина *ограничение* существует два омонима, имеющих различное звучание в английском языке: ограничения, которые нельзя нарушать – "constraint", и ограничения, которые свойственны сущностям – "restriction". Также определён термин *инвариант* как логическое выражение условия, истинность которого необходимо соблюдать.

Различия между данными, информацией и знаниями могут быть важными при выборе терминологии онтологического моделирования, концептуального моделирования баз данных или представления знаний. Термин *данные* используется для отдельных упорядоченных фактов, характеризующих объекты или процессы предметной области [16]. *Информацией* являются осмысленные или воспринимаемые данные [17]. *Знаниями* описываются закономерности предметной области (принципы, связи, законы, факты) [16].

### 7.3 Термины информационного поиска

Из области информационного поиска важно включить термины, характеризующие качество результатов работы алгоритмов поиска. В частности, методы информационного поиска активно используются для согласования онтологий. *Полнота выборки* – критерий оценки эффективности алгоритма по его способности находить все релевантные объекты, а *точность выборки* – находить только релевантные объекты. *Релевантность* – оценка того, насколько найденный объект или множество объектов соответствуют нуждам агента. *Сложность алгоритма* характеризуется соотношением между размером входных/выходных данных задачи и максимальным количеством элементарных операций (*временная сложность*), проделываемых алгоритмом для решения задачи указанного размера, либо между размером входных/выходных данных задачи и объёмом используемой алгоритмом памяти (*пространственная сложность*).

### 7.4 Термины информационных систем

В информационных системах и распределённых системах онтологии занимают заметное место, так как позволяют рассуждать о семантике данных.

*Информационной системой* называют комплекс средств для хранения информации и ее обработки, позволяющий отвечать на информационные запросы агентов. *Информационный ресурс* – совокупность данных, организованных для эффективного получения достоверной информации.

*Информационная коллекция* – систематическое собрание однородных объектов, представляющих научный, исторический, художественный или какой-либо другой интерес. *Электронная библиотека* – информационная система, представляющая доступ к коллекциям электронных документов, баз данных, сервисов, методов, программ и так далее, снабженная средствами навигации и поиска.

*Интероперабельность* – это способность систем работать совместно как одна интегрированная система. В глоссарии разграничены *технический*, *синтаксический* и *семантический* уровни интероперабельности. На онтологии часто возлагают задачи достижения последней.

## 7.5 Термины, употребляемые в лингвистике и семантике

Определены виды терминологической информации, эти термины принадлежат рубрике «Лингвистика». В лингвистике *термин* – это слово или словосочетание, которое используется в конкретном контексте с конкретным смыслом. *Терминология* включает набор терминов предметной области. Любые слова или устойчивые словосочетания являются *лексическими единицами*. *Вербальное определение* – установление смысла термина, формулируемое в языке с помощью слов или специальных знаков.

Термин *словарь* соответствует набору всех слов (терминов, символов) языка или коллекции документов, либо используемых агентом. Термин *контролируемый словарь* используется для ограниченной выборки терминов словаря. *Тезаурусом* называют словарь с дополнительной информацией о связях терминов, таких как синонимы, омонимы, родовидовые отношения, часть/целое, т. е. чаще всего используются некоторые виды семантических отношений. Термин *глоссарий* означает словарь (или контролируемый словарь) предметной области с определениями на естественном языке.

*Семантическими отношениями* являются отношения между двумя понятиями, терминами, лингвистическими единицами. В глоссарий включены следующие семантические отношения: *синонимия* или отношение *эквивалентности*, *омонимия*, *антонимия*, *ассоциативное отношение*, *родовидовое отношение* (выше/ниже, гипернимия/гипонимия, обобщение/специализация), отношение *часть/целое* (меронимия/холонимия), *отношение классификации* (класс/экземпляр), *отношение зависимости*.

## 7.6 Логические термины

Термины основ математической логики [18] включаются в глоссарий по онтологическому моделированию, так как онтологии описывают логические теории, относящиеся к понятиям предметной области. Большинство основных терминов математической логики постоянно используются в онтологическом моделировании. Глоссарий определяет такие термины, как *высказывание (пропозиция), высказывательная переменная, высказывательные связи и их виды (конъюнкция, дизъюнкция, отрицание, импликация, эквивалентность), кванторы, свободная и связанная переменные, формула, предикат, отношение, функция, терм и атом.*

С отношениями связаны их свойства: *арность, транзитивность, симметричность, рефлексивность, функциональность, обратные отношения, область определения и область значений.*

Перечислены виды логик, встречающиеся в онтологическом моделировании, включая следующие: *логика высказываний, логика предикатов первого порядка, дескриптивная логика* (в основном, пользуются этим переводом англоязычного термина *description logic*, однако точность перевода вызывает сомнения, и по поводу русскоязычного термина ведутся дискуссии, предлагающие такие варианты как *дескрипционная логика, логика описаний, модальная логика, немонотонная логика.*

*Аксиомы* предполагаются истинными без доказательства, *теоремы* – формулы, которые можно вывести из аксиом, *правила вывода* являются функциями из множества формул (*антецедента*) в формулу (*консеквент*). *Формальная система* – система, состоящая из формального языка и правил вывода. *Теория* – формальная система с языком 1-го порядка, набором аксиом (логических и нелогических) и правил вывода.

*Логическое определение* устанавливает смысл термина. *Аксиоматическое определение* формируется через множество аксиом (*постулатов*), которые ограничивают область толкования термина.

Термины, касающиеся различных задач логического вывода, используемых в онтологическом моделировании, и связанных с выводом свойств логических формул и теорий также необходимо включить в глоссарий. Для англоязычных терминов важно подобрать русскоязычные аналоги, которые традиционно используются в логике. Эти же термины будут использоваться и для задач логического вывода в области онтологического моделирования.

Программа, способная делать логические выводы из множества фактов и аксиом, называется *машиной вывода*. В основе вывода лежит *следование, импликация (entailment, implication)*. Многие логические задачи манипулируют понятиями интерпретации и модели. *Интерпретация* – в логике приписывание некоторого содержательного



смысла, значения символам и формулам формальной системы. *Моделью* (это омоним термина «модель» из области моделирования) является интерпретация, удовлетворяющая теории при всех значениях переменных.

*Доказуемость формулы* (provability) – существование известного набора правил, по которым строится доказательство формулы. *Разрешимость формулы* (decidability) – существование эффективного алгоритма определения, является ли формула теоремой.

*Правильность, валидность* (validity) – свойство формулы (или теории), которая истинна (а для теории – истинны её аксиомы) для любых интерпретаций. *Выполнимость, совместность* (satisfiability) – свойство формулы (теории), для которой существует хотя бы одна интерпретация, делающая формулу (аксиомы теории) истинной. В частности, этот термин используется в машинах вывода дескриптивной логики. *Непротиворечивость, целостность, согласованность* (consistency) – свойство теории, заключающееся в невыводимости из неё противоречия. *Противоречием* (contradiction) является пара высказываний, одно из которых является отрицанием другого. Для полных логик свойство выполнимости эквивалентно свойству непротиворечивости.

Корректность (soundness) – свойство логической системы, в которой правила вывода доказывают только правильные (валидные) формулы по отношению к её семантике. *Полнота* (completeness) – свойство аксиоматической теории, в которой все формулы, истинные при рассматриваемой интерпретации, доказуемы. Полная система содержит все возможные теоремы, не противоречащие интерпретации.

*Поглощение, включение* (subsumption) между свойствами имеет следующий смысл: свойство  $p$  поглощает свойство  $q$ , если и только если для любого состояния все экземпляры  $q$  являются также экземплярами  $p$ . В дескриптивной логике задача доказательства поглощения сводится к задаче доказательства выполнимости теории.

## 7.7 Разновидности деятельностей

В рубрике «Деятельность», в первую очередь, определены термины, фигурирующие в описаниях разного рода деятельностей. Это такие термины, как *действие, задача, предметная область, агент* (человек, группа людей или машина), *эксперт, экспертное решение*. Экспертов в области онтологического моделирования называют *онтологами* (ontologist). В рубрике определены деятельности, принадлежащие также концептуальному моделированию: *абстракция, классификация, категоризация, кластеризация, типизация*.

Все деятельности, производимые над онтологиями, можно разделить на деятельности по разработке онтологий, по их согласованию и по проверке их качества. Здесь важно подобрать подходящие русскоязычные

аналоги англоязычным терминам. Многие термины заимствованы из глоссария деятельностей проекта NeOn [19], однако необходимо заметить, что термины завязаны на конкретные особенности методологии создания и манипуляции онтологиями, используемой в проекте, и не все они могут быть использованы в исходном виде.

Деятельности, связанные с разработкой онтологий, представляются следующими терминами.

*Редактирование* (edition) онтологии – изменение онтологии посредством редактора. *Модификация* (modification) – законченное внесение изменения в онтологию. *Развитие* (evolution) – внесение модификаций в процессе жизненного цикла онтологии. Термин *модернизация* (upgrade) используется при замене существующей онтологии новой версией онтологии. С этим связана *поддержка версий* или *многоверсионности* онтологии (versioning). *Обновлением* (update) является внесение в онтологию незначительных изменений, которые не могут рассматриваться как модернизация. *Локализацией* (localization) называется адаптация онтологии к определённом естественному языку или культуре. *Подрезание* или *обрезка* (pruning) онтологии производится для удаления из неё знаний, более не актуальных.

Процесс *пополнения* (enrichment) онтологии заключается в дополнении онтологии новыми спецификациями и обобщает в себе: *специализацию* (specialization) – увеличение степени детализации понятий; *расширение* (extension) – увеличение охвата предметной области онтологии; *аннотирование* (annotation) – снабжение онтологии метаданными или комментариями.

*Сбор знаний* для онтологии (knowledge acquisition) включает деятельности по извлечению знаний из различных источников. Различают *извлечение* (elicitation) – работу по сбору знаний у экспертов предметной области, – и *обучение* онтологии (learning), опирающееся на автоматические или полуавтоматические методы анализа различных источников знаний. Необходимо заметить, что в сбор знаний входит и сбор индивидов, которые важны для определения понятий онтологии. В отличие от представлений проекта NeOn, заполнение экземплярами базы знаний, для которой онтология является схемой, частью процесса сбора знаний для онтологии не являются.

*Формализация* (formalization) онтологии – это процесс выражения концептуализации предметной области в соответствии с парадигмой представления знаний, предлагаемой языком моделирования; *реализация* (Implementation) онтологии – кодирование онтологии в соответствии с синтаксисом формального языка.

*Рейнжиниринг* (reengineering) онтологии предполагает извлечение знаний из онтологии для концептуализации предметной области и создание другой онтологии с использованием этих знаний. *Реструктуризация* (restructuring) онтологии – это исправление и

реорганизация знаний исходной концептуализации в результате анализа существующей онтологии и синтеза изменённой онтологии. *Трансляция* (translation) онтологии – преобразование онтологии с одного языка моделирования на другой.

*Спецификация* онтологии – процесс составления требований к онтологии либо сам набор требований, которым онтология должна соответствовать, таких как, целей создания онтологии, целевой группы, планируемого использования. *Документация* онтологии – процесс составления документов и пояснительных комментариев, порождаемых в течение всего процесса формирования онтологии, а также сам набор таких документов.

Термин *согласование онтологий* (reconciliation) [20] означает приведение в соответствие двух или более онтологий, обычно требующий изменений одной стороны или даже обеих сторон. Цели согласования онтологий различны, они перечислены ниже.

*Выравнивание* (alignment) онтологий заключается в установлении соответствий между двумя или более онтологиями и их хранении и использовании. *Соответствиями* могут быть отношения между элементами разных онтологий: понятиями, классами, экземплярами, свойствами или формулами. *Сопоставление* (matching) онтологий – вычисление соответствий между понятиями разных онтологий с применением лексических, структурных и других методов. Сопоставление онтологий можно рассматривать как вспомогательное средство, способствующее выравниванию онтологий. *Сравнение* (comparison) онтологий – нахождение различий между двумя или более онтологиями или между двумя или более модулями онтологии. *Отображение* (mapping) онтологий фактически является направленным вариантом выравнивания онтологий, оно заключается в процессе построения правил отображения понятий одной онтологии в понятия другой. Результирующее множество *правил отображения* (mapping rule) одинакового направления как разновидностей соответствий также называется «отображением». Элементы отображаемой онтологии могут присутствовать в нём только единожды. Отметим, что проект NeOn некорректно использует термин «отображение онтологий», приравнивая его к термину «выравнивание онтологий».

Термин *слияние онтологий* (ontology merging) применяется для создания новой онтологии из двух или более исходных онтологий или модулей, возможно, пересекающихся. Исходные онтологии при этом остаются неизменными.

Термином *интеграция онтологий* (integration) следует пользоваться в том случае, если одна онтология (вероятно, описывающая часть предметной области) включается в другую онтологию. Также в целевую онтологию включаются утверждения, соединяющие эти онтологии. Они называются *аксиомами сочленения* (bridge axiom, articulation axiom).

*Гармонизация* (harmonization) онтологий не является распространённым термином. Он применялся как интеллектуальный процесс связывания онтологий на уровне концептуализаций в процессе дискуссий [21].

Деятельности, отвечающие за проверку качества онтологий, описываются следующими терминами.

В русском языке используются два омонима *оценка онтологии* для англоязычных терминов, означающих проверку качества онтологии относительно требований пользователя (ontology assessment) и относительно технических критериев (ontology evaluation). Последний термин обобщает термины *валидация онтологии* (ontology validation), означающий проверку на модели мира, представляющей концептуализацию, и *верификация онтологии* (ontology verification), означающий оценку соответствия онтологии спецификации требований. *Диагностика* (diagnosis) онтологии инициируется в процессе валидации онтологии и выявляет части онтологии, определяющие её некорректность и неполноту. *Исправление* (repair) онтологии – это исправление выявленной некорректности или неполноты.

## 8 Акценты усилий и перспективы развития глоссария

Глоссарий по онтологическому моделированию в сегодняшнем его состоянии нуждается в согласовании и развитии. Предпочтительным является его коллективное обсуждение и изменение специалистами в области онтологического моделирования в соответствии с разработанными и представленными в данной работе принципами.

Особое внимание при изменении и развитии глоссария сообществом специалистов необходимо уделять следующим вопросам, не имеющим тривиального решения:

- обсуждение разумного охвата глоссарием областей, смежных с онтологическим моделированием и несмежных, но необходимых в глоссарии для разграничения терминологий разных областей знания;
- выбор и согласование русскоязычной терминологии и контекстов её применения с учётом исторического применения терминов в смежных областях;
- поиск представительных и изначальных источников определений терминов, устранение определений терминов, данных в контекстах производных предметных областей;
- оценка качества определений и взаимное их согласование.

Для оценки качества определений и их взаимного согласования возможно применение методов морфологического и семантического анализа определений терминологических словарей [22].

Выбор источников определений усложняется тем, что часто даже в фундаментальных статьях, вводящих термин, нет чётких определений. В этом случае могут быть полезны глоссарии разных областей знания [19, 18, 23-25]. Однако определения в них необходимо рассматривать критически, так как они бывают адаптированы к задачам и предметным областям авторов. Необходимо отметить, что в открытом для редактирования проекте энциклопедии Wikipedia [26] качество описаний многих терминов низкое, они подчас слабо относятся к заявленной теме.

## 9 Заключение

Для создания русскоязычного глоссария по онтологическому моделированию были разработаны и представлены структуры статей глоссария, принципы и средства его коллективного обсуждения и развития. На момент написания статьи глоссарий содержит около 200 терминов с определениями, расположен на общедоступном wiki-ресурсе и готов к формированию рабочей группы и коллективному обсуждению. Эту работу желательно соотносить с работой над созданием учебных программ в области онтологического моделирования. Результаты работы могут послужить согласованию русскоязычной терминологии в области онтологического моделирования, сближению точек зрения на предметную область онтологического моделирования и на смежные с нею области различных сообществ. После формирования и согласования русскоязычного глоссария с англоязычными эквивалентами было бы естественным довести глоссарий до полноценного двуязычного варианта.

## Литература

1. Труды Симпозиума «Онтологическое моделирование», г. Звенигород, 19 – 20 мая 2008 г., ред. Л. А. Калиниченко – М: ИПИ РАН, 2008. – 303 с. – ISBN 978-5-902030-54-6
2. OWL Web Ontology Language guide. W3C recommendation, 10 February 2004. – URL: <http://www.w3.org/TR/2004/REC-owl-guide-20040210/>
3. Рабочая группа симпозиума «Онтологическое моделирование». – URL: <http://ontology.ipi.ac.ru/>
4. *M. del C. Suárez-Figueroa, A. Gómez-Pérez.* Towards a glossary of activities in the ontology engineering field. – LREC. – Marrakech, 2008.
5. Cicero. – URL: [http://cicero.uni-koblenz.de/wiki/index.php/Main\\_Page](http://cicero.uni-koblenz.de/wiki/index.php/Main_Page)
6. *T. R. Gruber.* A translation approach to portable ontologies. // Knowledge Acquisition. – Vol. 5 (2). – 1993. – URL: <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>
7. *N. Guarino.* Formal ontology and information systems. // Proc. of FOIS'98, Trento, Italy. – Amsterdam: IOS Press, 1998. – P. 3-15.

8. A. B. Markman. Conceptual representations in psychology. // Encyclopedia of Cognitive Science. L. Nadel (ed.). – London: Macmillan, 2003. – Vol. 1. – P. 670-673
9. M. R. Genesereth, N. J. Nilsson. Logical foundation of artificial intelligence. – Los Altos, California: Morgan Kaufmann, 1987.
10. W. O. Quine. From a logical point of view // Nine Logico-Philosophical Essays. – Harvard University Press, Cambridge, Mass, 1961.
11. A. Newell. The knowledge level. // Artificial Intelligence. – Vol. 18 (1). – 1982.
12. И. Л. Артемьева. Онтологии предметных областей и их использование при создании программных систем. Труды Симпозиума «Онтологическое моделирование», г. Звенигород, 19-20 мая 2008 г., ред. Л. А. Калиниченко. – М: ИПИ РАН, 2008. – С. 83-113 – ISBN 978-5-902030-54-6
13. М. Р. Коголовский, Л. А. Калиниченко. Концептуальное моделирование в технологиях баз данных и онтологические модели. Труды Симпозиума «Онтологическое моделирование», г. Звенигород, 19-20 мая 2008 г., ред. Л. А. Калиниченко. – М: ИПИ РАН, 2008. – С. 114-148 – ISBN 978-5-902030-54-6
14. B. Hjørland. Metadata. Lifeboat for knowledge organization. – URL: [http://www.iva.dk/bh/lifeboat\\_ko/concepts/metadata.htm](http://www.iva.dk/bh/lifeboat_ko/concepts/metadata.htm)
15. N. Guarino, L. Schneider. Ontology-driven conceptual modelling. // Proc. of the 21st International Conference on Conceptual Modeling, LNCS. – Vol. 2503. – Springer-Verlag, 2002
16. Т. А. Гаврилова. Онтологическое мышление и визуальные модели в инженерии знаний. Труды Симпозиума «Онтологическое моделирование», г. Звенигород, 19-20 мая 2008 г., ред. Л. А. Калиниченко – М: ИПИ РАН, 2008. – С. 46-70 – ISBN 978-5-902030-54-6
17. К.-Е. Sveiby. What is information?, 1994. – URL: <http://www.sveiby.com/articles/Information.html>
18. P. Suber. Glossary of first-order logic. – URL: <http://www.earlham.edu/~peters/courses/logsys/glossary.htm>
19. A. Gómez-Pérez. NeOn glossary of activities. – URL: <http://kmi.open.ac.uk/events/sssw08/presentations/Gomez%20Perez-neonglossaryofactivities.pdf>
20. J. Euzenat, P. Shvaiko. Ontology matching. – Springer-Verlag, 2007
21. S. Mazzocchi. The psychology of ontology harmonization. – Stefano's Linotype, 2004. – URL: <http://www.betaversion.org/~stefano/linotype/news/60/>
22. К. К. Боярский, Е. А. Каневский, Г. В. Лезин и др. Автоматизация процесса извлечения онтологической информации из вербальных терминологических словарей (на примере терминологического словаря задачи межзвездного поглощения). – RCDL'2010. – Казань, 2010
23. Lifeboat for knowledge organization. – URL: [http://www.db.dk/bh/lifeboat\\_ko/home.htm](http://www.db.dk/bh/lifeboat_ko/home.htm)
24. R. Šumrada. The internal and external views of information system. // COST G9, Bremen. – 2001.
25. G. Booch. Object-oriented analysis and design with applications. – Addison-Wesley, 2007. – ISBN 0-201-89551-X.
26. WikiPedia, the free encyclopedia. – URL: [http://en.wikipedia.org/wiki/Main\\_Page](http://en.wikipedia.org/wiki/Main_Page)