

## OWL КАК СТАНДАРТНАЯ МОДЕЛЬ ПРЕДСТАВЛЕНИЯ ТРАНСДИСЦИПЛИНАРНЫХ ЗНАНИЙ В SEMANTIC WEB

Андрей Михайлюк

**Аннотация:** *Рассматриваются причины появления OWL как стандарта для представления онтологий в Semantic Web, его структура и возможности для представления онтологий трансдисциплинарных знаний. Предложены рекомендации по использованию OWL на междисциплинарном уровне.*

**Ключевые слова:** *OWL, междисциплинарные исследования, онтология.*

**ACM Classification Keywords:** *1.2.4. Knowledge Representation Formalisms and Methods*

---

### Введение

Не вызывает сомнений тот факт, что подавляющее большинство знаний, которыми владеет человечество, представлено в цифровом виде в сети Интернет. Простота доступа, надежность хранения и широкий диапазон вариантов представления (включая интерактивные) делают Интернет неотъемлемым источником информации в любых научных исследованиях. Особенно важным данный вид представления и доступа к знаниям выступает в междисциплинарных исследованиях, где рабочие группы территориально и инфраструктурно отдалены. Однако, невзирая на объективные преимущества цифрового представления знания в компьютерных сетях для трансдисциплинарных исследований, данная область требует дальнейшего усовершенствования.

---

### Постановка задачи

Одним из перспективных направлений исследования являются исследования моделей представления онтологий, предназначенных для использования в различных областях науки и хозяйства, в частности, модель и язык OWL, рекомендованный консорциумом W3C и де-факто ставший стандартом представления онтологий в развивающемся пространстве Semantic Web. Несмотря на то, что назначением OWL является представление онтологических знаний в глобальном информационном пространстве, текущие исследования, исходя из существующих онтологий, онтолого-ориентированных проектов и систем, ограничиваются узкоспециализированными областями: медицина, биология, незначительная часть социальных и технических наук [W3C Wiki, Semantic Web Wiki]. Несомненно, внимания требует **слабо исследованная область применения OWL на трансдисциплинарном уровне**, что в перспективе послужит интегрирующим звеном в задаче построения Semantic Web. Целью данной работы является анализ структуры и возможностей OWL в рамках применения на междисциплинарном уровне и **выработка рекомендаций для представления в OWL трансдисциплинарных онтологических знаний**.

---

---

## Предпосылки создания OWL

---

Первым шагом в создании глобальной межпредметной коммуникации, которая на порядок повысила эффективность и возможности взаимодействия исследователей, стало развитие Интернет и его основного способа представления информации (в т.ч. и знаний) – языка гипертекстовой разметки **HTML**. Основной функцией HTML является **оформление** текстовой и графической информации в структурированном виде (абзацы, списки, таблицы и т.п.). Его структура призвана упростить и унифицировать **способ визуального отображения** информации в окне браузера. Данные возможности позволяют авторам наглядно представлять результаты своих исследований, соискателям – легко понимать содержание изложенной информации, а компьютерным системам – упростить алгоритмы ее отображения.

Следующим вопросом, возникшим с ростом глобальной сети, стал поиск информации. Очевидно, что исследователь при поиске необходимой информации не в состоянии пересмотреть (а тем более понять) содержание миллионов источников. Для этого были созданы **автоматические системы поиска** ресурсов, релевантных запросам пользователя, а язык HTML был дополнен **метаинформацией** (тегами), **идентифицирующей общее содержание** или область, к которой относится ресурс. Отметим, что такая метаинформация уже описывает не способ отображения основной информации ресурса, а его **общую семантику**. Поисковые системы значительно упрощают поиск ресурса, необходимого исследователю. Однако с дальнейшим ростом количества информации в глобальной сети отфильтрованный список возможных релевантных ресурсов сам стал настолько объемным, что обработка исследователем такого количества источников с целью выделения знаний из информационных ресурсов стала невозможной. Эволюция поисковых систем лишь количественно решала данную проблему, сужая круг поиска, в то время как выделение знаний все еще оставалось исключительно задачей человека. Кроме того, подход обособленной идентификации общей семантики ресурса по специальным тегам не дает гарантии, что ресурс действительно отвечает заявленной семантике. Так, множество ресурсов для поднятия собственного рейтинга в поисковых системах помечаются наиболее популярными тегами вместо наиболее релевантных содержанию ресурса.

С целью преодоления указанных проблем возникло два основных направления: попытка **автоматического выделения семантики из самого ресурса** (а не из искусственно прописанных тегов) и **внесение в структуру ресурса семантической составляющей** наряду (а иногда и вместо) визуальной разметки. Первый путь включает обработку естественного языка, распознавание графических образов, аудио-сигналов и т.п. Данный подход дает возможность получать знания из уже созданных и значительно упрощает выделение знаний из новых ресурсов, но реализация такого подхода на должном уровне выступает научно-технической проблемой, надежные способы решения которой не определены и по сегодня. Следует, однако, отметить существование методов и систем, которые реализуют этот подход поверхностно (как правило, средствами математической статистики) но достаточном для ряда прикладных задач уровне (реферирование естественно-языкового текста, поиск людей по фотографии или фотороботу и т.д.). Другой подход заключается в создании информационных ресурсов, структура которых изначально описывала бы семантику информационных составляющих, таких как термины,

определения, графические изображения и др. Такая метаинформация должна быть простой для обработки существующими инструментами работы с Web-ресурсами в автоматическом режиме. Так, на основе языка XML как расширения HTML был предложен **фреймворк описания ресурсов RDF** (Resource Description Framework) и определена **схема описания RDFS** (RDF Schema). Концептуально RDF всего лишь определяет широко известный принцип формирования фактов при помощи бинарных отношений между двумя ресурсами [Hoekstra 2009]. В дополнение, RDFS вносит семантические структуры, такие как таксономические отношения, возможность типизации (определение некоторых ресурсов как типов для других ресурсов), определения классов ресурсов, их свойств и т.п. RDFS как модель является достаточно абстрактной для выражения любых фактов, включая правила, ограничения и мета-утверждения. Однако, такая абстрактность вынуждает разработчиков онтологий вводить собственную мета-семантику и семантику верхнего уровня в каждую онтологию, что значительно снижает степень определенности при анализе RDF/RDFS онтологии автоматическими машинами вывода. Очевидной стала необходимость внесения общепринятой семантики (логические и множественные отношения, транзитивные, симметричные и др. свойства отношений, арифметические и списковые примитивы, эксплицитные понятия тождественной истинности и ложности и др.) в спецификацию языка представления онтологий.

Внесением общей онтологической семантики в модель представления знаний одновременно занимались две организации: DARPA (US Defense Advanced Research Projects Agency) и European Research Area в рамках проекта Information Society Technologies. Независимо были разработаны две соответствующие модели – **DAML** (DARPA Agent Markup Language) [DARPA, McIlraith, 2001] и **OIL** (Ontology Inference Layer или Ontology Interchange Language) [Fensel, 2001]. Обе базируются на RDF/RDFS и в значительной мере пересекаются. В дальнейшем две модели были объединены в одну **DAML+OIL** [Connolly, 2001]. Как отмечается в [Hoekstra 2009], она получила семантику дескриптивной логики от OIL, что дало возможность построения эффективных машин вывода, и совместимость с RDF/RDFS от DAML, что позволило использовать DAML+OIL в рамках Semantic Web. Приняв во внимание все преимущества DAML+OIL, концерн W3C создал аналогичную модель и язык с большей выразительной мощностью и точностью описания онтологий – **OWL** (Ontology Web Language). Последним на сегодня является стандарт **OWL2** [OWL2 Overview, 2012] – в дальнейшем изложении, если иное не указано явно, будем использовать аббревиатуру “OWL” в значении последнего доступного стандарта OWL, то есть OWL2.

---

## Язык и модель OWL

---

Язык OWL можно рассматривать в разных аспектах. Так, “горизонтально” можно выделить **синтаксис** и **семантику** языка OWL, а также **структуру** OWL онтологии. Семантика и абстрактная структура онтологии определяют **модель OWL**.

### Синтаксисы и структура OWL

Единственным необходимым для OWL **синтаксисом** является **RDF/XML**. Он гарантирует поддержку обмена онтологиями между гетерогенными системами, которые могут даже не поддерживать сам язык OWL как таковой (чистые RDF/RDFS системы) – сериализация в данный синтаксис определяет правила

приведения элементов языка OWL к абстрактным конструкциям RDF, записанным в виде иерархии XML тегов.

Другим синтаксисом, который использует XML как низкоуровневый язык структуризации элементов OWL онтологии является синтаксис **OWL/XML**. Его отличие от RDF/XML заключается в том, что сериализатор не описывает встроенные элементы языка OWL в терминах RDF, полагаясь на то, что система-получатель должна корректно распознать и проинтерпретировать специфические элементы OWL. Целью введения данного синтаксиса является предоставление возможности оптимального представления онтологий для систем, ориентированных на OWL как внутренний язык и XML как формат обмена.

Важным, но не обязательным для поддержки, является **Functional Syntax**. Его особенность заключается в компактной структурной записи элементов OWL в виде, подобном записи предикатов математической логики или функциональным языкам (например LISP) – отсюда и возникло название синтаксиса. Как отмечено в [OWL2 Overview, 2012], данный синтаксис наилучшим образом отображает структуру онтологии. Стандарт [OWL2 Structure, 2012], который описывает структуру абстрактной онтологии, использует именно Functional Syntax.

**Manchester Syntax** также оперирует встроенными элементами OWL, которые должны распознаваться системой, и направлен на представление, оптимизированное под сохранение и загрузку онтологий, базирующихся на дискриптивной логике.

Последний опциональный синтаксис OWL – это **Turtle** (Terse RDF Triple Language) – компактный способ записи графов бинарных отношений RDF.

На общем уровне OWL как языка стандарт [OWL2 Structure, 2012] выделяет три составляющих:

- **Сущности**, такие как классы, свойства и экземпляры;
- **Выражения** – сложные записи в описываемом домене, семантика которых соответствует множествам определенных анонимных сущностей;
- **Аксиомы** – утверждения, которые считаются тождественно истинными в универсуме дискурса.

### Уровни использования OWL

Для целей практического использования OWL, целесообразно разделять его элементы также по уровню использования:

- **Мета-уровень** онтологии, где определяется, из каких частей состоит онтология (импорт других онтологий), какие словари доступны в онтологии, какие сокращения будут использоваться в онтологии, название, версия, аннотация к онтологии и т.п. Структуры и элементы данного уровня значительно расширяют возможности удаленного и разделяемого пользования онтологиями, их классификации, обмена, оптимизации и т.д.;
- **Уровень структур данных** – часть словаря и правил OWL, которые определяют доступные типы литералов (например, текстовые строки, целые числа, даты и время), уникальные и универсальные способы идентификации сущностей (IRI [Duerst, 2005], Unicode [Yergeau, 2003]), методы обращения к бинарным ресурсам и др. Данный уровень обеспечивает максимальную универсальность относительно локации и области использования OWL онтологий;

- **Онтологический уровень** – структуры и элементы OWL, которые используются для непосредственного описания содержательной части онтологии. Составляющие данного уровня определяют выразительную мощь, семантику, противоречивость и возможности анализа OWL онтологий.

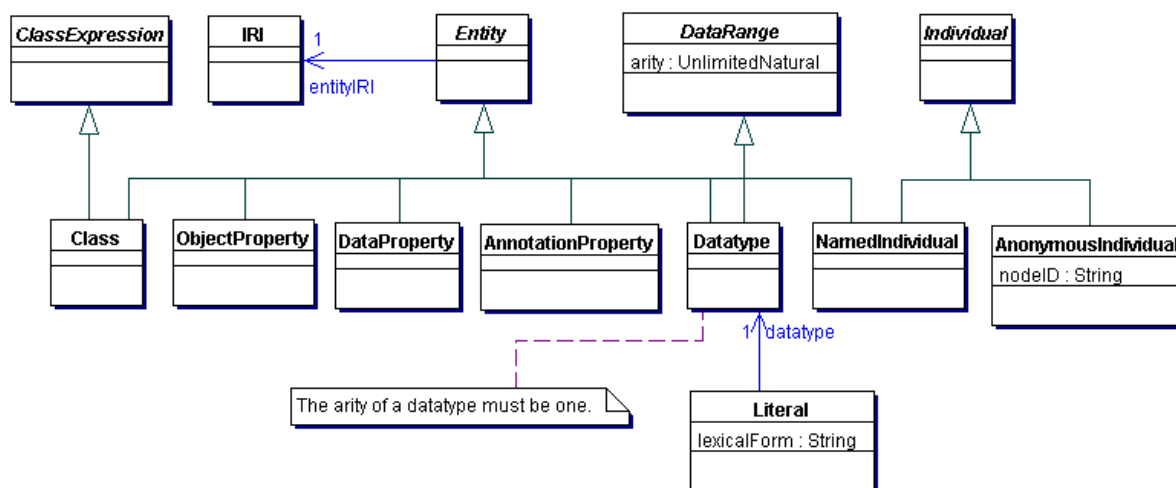


Рис. 1. Сущности OWL и их зависимости

Языковые составляющие по уровням использования OWL взаимно пересекаются, то есть на каждом из трех уровней OWL предоставляет набор языковых средств всех трех типов. Поскольку в цели данной работы входит анализ прежде всего онтологических (а не технических) возможностей OWL, рассмотрим лишь последний уровень использования OWL – онтологический уровень.

Базовыми элементами OWL являются **сущности**. Общую структуру зависимостей между типами сущностей и их место в OWL отражает рис. 1 [OWL2 Structure, 2012].

Сущности онтологического уровня OWL включают:

- **Классы (Class)**. Общепринятое понятие в онтологическом инжиниринге, которое описывает множество элементов (объектов, действий, явлений) объективного мира, собранных по наличию некоторого общего признака. Другие источники могут также использовать аналогичные термины “понятие”, “фрейм” и др. Аналогично многим другим онтологическим высокоуровневым языкам [Genesereth, Sowa, 2000 (1)], OWL содержит как минимум два специальных имплицитно определенных класса: *owl:Thing* и *owl:Nothing*;
- **Типы данных (Datatype)**. Определяет естественное (качественное) ограничение на множество литеральных значений и их специфическую семантику (например, тип данных *integer* ограничивает связанную переменную множеством целых чисел  $Z$  и определяет набор математических операций над такой переменной как множество математических операций над полем  $Z$ );
- **Объектные свойства (ObjectProperty)** описывают бинарные отношения между экземплярами классов. В отличие от широко известного разделения характеристик объектов на отношения (N-

арные отношения) и атрибуты (унарные отношения), в OWL свойством объекта называется исключительно бинарное отношение. В соответствии с отображением на семантику RDF [Patel-Schneider, 2012], объектные отношения OWL являются RDF классами, а поэтому могут формировать иерархии, иметь экземпляры и т.п.;

- **Свойства данных (*DataProperty*)** – аналогично объектным свойствам характеризуют (вводят ограничения на) экземпляр класса путем определения бинарного отношения, однако вторым операндом здесь выступает не другой экземпляр, а литерал. Свойства данных, как и объектные свойства, также являются RDF классами со всеми вытекающими возможностями;
- **Экземпляры классов (*Individual*)**, которые описывают элементы объективного мира, типом которых являются некоторые классы. Важным отличием экземпляров от всех других сущностей является то, что их имена локальны для данной онтологии, то есть если имена сущностей в данной и импортированной онтологиях совпадают, такие сущности считаются различными.

**Выражения OWL** подразделяются на выражения свойств (объектных и свойств данных) и выражения классов. **Выражения объектных свойств**, согласно [OWL2 Structure, 2012] (см. рис. 2), помимо самих свойств, может описывать свойство, обратное данному. Данная часть OWL, несмотря на столь незначительную выразительность, на данный момент не нуждается в расширении, т.к. основные описательные возможности в отношении свойств обеспечиваются многообразием выражений классов (т.к. любые свойства OWL сами являются классами).

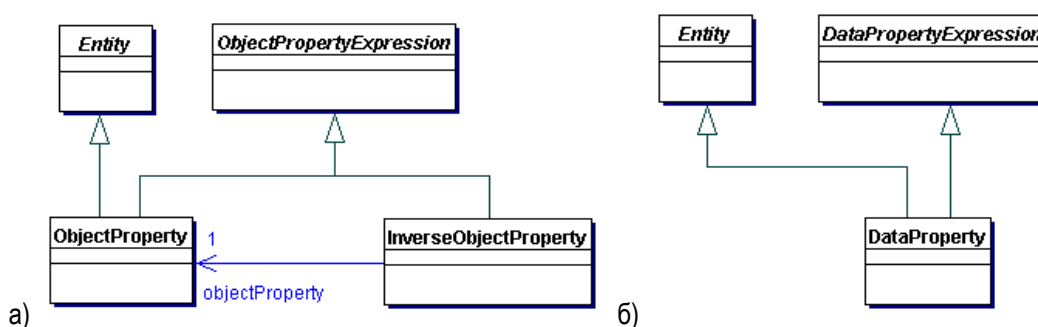


Рис. 2. Структура выражений свойств OWL:

а) выражения объектных свойств; б) выражения свойств данных.

**Выражения классов**, как говорилось ранее, имплицитно определяют новый класс, удовлетворяющий указанным характеристикам. Такими характеристиками и соответствующими типами выражений являются:

- **Теоретико-множественные операции**, соответствующие основным способам задания множеств, а именно **пересечение (*ObjectIntersectionOf*)**, **объединение (*ObjectUnionOf*)**, **дополнение (*ObjectComplementOf*)** и **перечисление элементов множества (*ObjectOneOf*)**. В терминах понятий как аналогов OWL классов, данная группа выражений описывает новое понятие через его объем, или, иначе говоря, количественно (экстенционально) определяет класс;

- **Ограничения по объектным свойствам**, в отличие от теоретико-множественных операций, определяют класс качественно, интенционально. Семантика этого типа выражений заключается в определении класса, экземпляры которого неким образом связаны заданным бинарным отношением (объектным свойством) с указанными экземплярами. Объектное свойство и экземпляры, формирующие его диапазон значений, так же могут быть заданы эксплицитно. Типы ограничения включают ограничения с **квантором существования** (*ObjectSomeValuesFrom*), **универсально квантифицированные** (*ObjectAllValuesFrom*), **индивидуальные** ограничения (*ObjectHasValue*), **самоограничения** (*ObjectHasSelf*), ограничения с **минимальной** (*ObjectMinCardinality*), **максимальной** (*ObjectMaxCardinality*) и **точной** (*ObjectExactCardinality*) **кардинальностью**. Отметим, что подобное многообразие вариантов интенционального определения классов наилучшим образом соотносится с применением OWL к представлению онтологий в междисциплинарных исследованиях, т.к. данный способ определения понятий является, с одной стороны, одним из основных в научном мире, а с другой – описывает определяемое понятие безотносительно к конкретным, специфическим для данной дисциплины, частным его (понятия) проявлениям или экземплярам;
- **Ограничения по свойствам данных** аналогичны предыдущему типу выражений, с тем лишь основным отличием, что операндом свойства является ограничение экземпляром данных (открытым или закрытым диапазоном, множеством, единичным или анонимным значением). Т.к. данный тип выражений, как и другие, определяет класс, а ограничивающим операндом свойства является не экземпляр класса, в нем отсутствует самоограничение, доступное в выражениях ограничения по объектным свойствам.

**Аксиомы** можно назвать наиболее важными элементами OWL, т.к. их многообразие во многом определяет богатство языка и его удобство в составлении онтологий, что особенно важно в описании сложных онтологий, представляющих домены научных дисциплин. С другой стороны, определение закономерностей в описываемом онтологией универсуме при помощи заранее предопределенного набора аксиом упрощает разработку машин вывода и других систем анализа онтологий, т.к. семантика каждой предопределенной аксиомы известна, а способ идентификации в онтологии сводится к сравнению с шаблонной строкой. В противовес этому, анализаторы “чистых” RDF онтологий должны определять, например, свойство транзитивности некоего отношения анализом структуры данного отношения, его родовых отношений и логических отношений, в которые данное входит как операнд. Как и выражения, аксиомы целесообразно разделять по типу описываемых сущностей:

- **Аксиомы выражения классов** определяют взаимоотношения между OWL классами, тождественно истинные в данной онтологии. Сюда входят отношения **родовидовой зависимости** (*SubClassOf*), **эквивалентный класс** (*EquivalentClasses*), **несовместный** (непересекающийся, отличный, дизъюнктивный) класс (*DisjointClasses*), **объединение несовместных классов** (*DisjointUnion*). Данные аксиомы являются мощным инструментом в определении родовидовой иерархии онтологии;

- **Аксиомы объектных свойств** можно подразделить на две группы согласно двойственности природы самих OWL свойств:
  - Ввиду того, что объектные свойства являются классами, над ними определен ряд аксиом, аналогичных аксиомам выражения классов. Это аксиомы утверждения **родовидовой зависимости** (*SubObjectPropertyOf*, *ObjectPropertyChain*), **эквивалентности** (*EquivalentObjectProperties*), **несовместности** свойств (*DisjointObjectProperties*) и, согласно рис. 2 (а), аксиомы утверждения **инверсности** двух объектных свойств (*InverseObjectProperties*);
  - С другой стороны, свойства в OWL являются бинарными отношениями, поэтому вторым подмножеством аксиом над ними являются утверждения о типе бинарного отношения, а именно **функциональность** (*FunctionalObjectProperty*) и **обратная функциональность** (*InverseFunctionalObjectProperty*), **рефлексивность** (*ReflexiveObjectProperty*) и **иррефлексивность** (*IrreflexiveObjectProperty*), **симметричность** (*SymmetricObjectProperty*) и **асимметричность** (*AsymmetricObjectProperty*), **транзитивность** (*TransitiveObjectProperty*). Так же, будучи бинарным отношением, объектное свойство имеет **область определения** и **область значения**, что так же может быть описано соответствующими аксиомами (*ObjectPropertyDomain* и *ObjectPropertyRange* соответственно).
- **Аксиомы свойств данных** включают утверждения, аналогичные аксиомам объектных свойств (*SubDataPropertyOf*, *EquivalentDataProperties*, *DisjointDataProperties*, *FunctionalDataProperty*, *DataPropertyDomain*, *DataPropertyRange*), за тем лишь исключением, что имея область определения экземпляра класса, а область значения – литеральный тип, не могут быть обратно-функциональными, рефлексивными и иррефлексивными, симметричными и асимметричными, транзитивными;
- **Определения типов данных** (*DatatypeDefinition*) служит инструментом для введения специфических (пользовательских) типов в онтологию ограничением существующих. Такие аксиомы, помимо удобства использования специфического для данной предметной области типа и автоматической поддержки операций над ним (унаследованной от родительского предустановленного типа), дает возможность анализировать специфические типы различных предметных областей в междисциплинарных исследованиях при интеграции онтологий с пользовательскими типами. Иначе говоря, пользовательские типы – это не только удобный технический прием, но и элемент семантики предметной области;
- **Определение ключей** (*HasKey*) помечает подмножество свойств экземпляров некоего класса как уникально идентифицирующих каждый экземпляр. Как и в случае с определением пользовательских типов, данная возможность, помимо очевидного влияния на логический вывод над онтологией, вносит семантику, специфическую для данной предметной области, что также может быть важно при анализе онтологий в рамках междисциплинарных исследований;



- **Утверждения** об экземплярах классов (также известные как **суждения** или **факты**) определяют **эквивалентность** (*SameIndividual*) или **различность** (*DifferentIndividuals*) каждого из перечисленных экземпляров, **принадлежность экземпляра к некому классу** (*ClassAssertion*), **присутствие** или **отсутствие свойства** у данного экземпляра (*ObjectPropertyAssertion* и *NegativeObjectPropertyAssertion*, *DataPropertyAssertion* и *NegativeDataPropertyAssertion*).

Следует также отметить, что такой богатый набор всевозможных аксиом и возможность явного указания истинности или ложности некоторого выражения в значительной степени упрощает создание онтологии верхнего уровня – основы для построения общей междисциплинарной онтологии из онтологий различных предметных областей.

---

### Использование OWL в онтологиях трансдисциплинарного уровня

---

Основной особенностью междисциплинарных исследований, по сравнению с исследованиями в рамках некой отдельной предметной области, является очерчивание нерешенных проблем, слабо исследованных вопросов и поиск новых знаний, находящихся на стыке дисциплин. Говоря об онтологическом инжиниринге, где каждая предметная область, ее знания и задачи описываются онтологиями, можно заметить, что основной задачей при проведении междисциплинарных исследований является **обобщение** онтологий. Обобщение включает в себя две основных составляющих: **построение онтологии верхнего уровня**, определяющей общенаучную, или хотя бы общую для данных двух онтологию, а так же эффективные механизмы непротиворечивой **интеграции онтологий** разных предметных областей.

Обсуждения подходов к **построению онтологий верхнего уровня** неоднократно проводились в многочисленных работах [Палагин, 2006, Найханова, 2005, Schneider, 2003]. В рамках данной статьи важно отметить, что OWL предоставляет широкий спектр возможностей для построения верхнеуровневых онтологий. Так, помимо общеиспользуемых аксиом описания таксономических иерархий и раскрытия содержания понятий категориального уровня, **особо полезными элементами OWL** в данном случае являются:

- Аксиомы, устанавливающие высокоуровневые **отношения эквивалентности** между классами (*EquivalentClasses*, *DisjointClasses* и *DisjointUnion*), позволяющие на категориальном уровне определить как точки соприкосновения различных предметных областей или теорий (что необходимо при интеграции доменных онтологий), так и потенциальные конфликты (например, при анализе онтологий из разных дисциплин легко установить недостаточность или противоречивость некоторых их частей посредством проверки на несовместность надклассов категориального уровня);
- Аксиомы объектных свойств, предназначенные для **построения иерархий отношений**, что является важной частью онтологической картины мира [Sowa, 2000 (2)]. К ним относятся аксиомы *SubObjectPropertyOf*, *ObjectPropertyChain*, *EquivalentObjectProperties*, *DisjointObjectProperties* и *InverseObjectProperties* и аналогичные отношения над литеральными типами;

- Аксиомы типизации объектных свойств как бинарных отношений, что позволяет на верхнем уровне определить **схему взаимозависимости неизвестных** на данный момент **классов и отношений** конкретных предметных областей приписыванием верхнеуровневому отношению некоего типа. Это аксиомы *FunctionalObjectProperty*, *InverseFunctionalObjectProperty*, *ReflexiveObjectProperty*, *IrreflexiveObjectProperty*, *SymmetricObjectProperty*, *AsymmetricObjectProperty*, *TransitiveObjectProperty*, *ObjectPropertyDomain* и *ObjectPropertyRange*;
- Возможность оперирования **анонимными сущностями** и введения **мета-описаний** позволяют более лаконично и абстрактно устанавливать категориальные правила и основные объективные законы (например, в OWL возможно определить понятия “тяжелый” и “легкий”, оперируя анонимными массами анонимных предметов, имеющими данные характеристики).

Для задачи **интеграции онтологий**, принадлежащих разным доменам знаний, основными рекомендациями являются рекомендации (а в перспективе – методики) по созданию OWL онтологий, пригодных для автоматического анализа на междисциплинарном уровне. Одной из очевидных рекомендаций в данном контексте является искусственное ограничение используемого в предметной онтологии синтаксиса OWL. Данная идея, в общем, не является оригинальной. Как известно, возможность использования логических и мета-онтологических конструкций на одном уровне приводит к неразрешимости OWL в общем смысле [Motik, 2007], а богатая выразительность значительно усложняет разработку эффективных алгоритмов анализа онтологий. Поэтому, наряду с полным языком OWL (который иногда называют OWL Full [Hoekstra 2009]), стандартом [Motik, 2012] определено также три его подмножества (профиля): OWL EL, OWL QL и OWL RL, предназначенных соответственно для оптимизированного вывода за полиномиальное время на онтологиях значительного объема; ускорения выполнения конъюнктивных запросов (в частности, на SQL-подобных языках) к базам со значительным количеством сравнительно небольших онтологий; оптимизации выполнения (вплоть до полиномиальной сложности) вывода на правило-ориентированных системах анализа RDF-подобных графов. Детальное рассмотрение указанных подмножеств выходит за рамки данной работы, поэтому, приняв искусственное ограничение синтаксиса языка за широко используемую практику, можно предложить следующие рекомендации:

1. Использование **интенциональных определений** понятий предметной области. Относительно OWL, это означает преимущественный способ введения классов через **наследование и ограничение по объектным свойствам**. Так, для описания нового класса, отличающегося от надкласса некоторым набором отношений (собственное содержание понятия), следует ввести в онтологию две записи: (1) аксиому родовидовой зависимости *SubClassOf* от надкласса и (2) аксиому родовидовой зависимости от анонимного класса (или классов), описанного выражением ограничения по объектному свойству *ObjectSomeValuesFrom* – от этого анонимного класса будут унаследованы новые свойства;
2. Максимальное описание **содержания понятий**, специфических данной дисциплине. Данная необходимость мотивирована тем, что формальное обобщение на междисциплинарном уровне возможно через имена понятий, либо через их содержание [Палагин, 2010]. Известными

---

проблемами инженерии знаний, особо остро стоящей в междисциплинарных исследованиях, являются синонимия и омонимия понятий. Поэтому, для качественного анализа онтологии некой предметной области каждое понятие рекомендовано раскрывать содержательно. Со стороны OWL, это должно выражаться в явном описании максимального количества объектных свойств классов и их значимых свойств данных. Как говорилось ранее, для описания объектных свойств класса необходимо ввести в онтологию аксиому родовидовой зависимости от анонимного класса (или классов), описанного выражением ограничения по объектному свойству *ObjectSomeValuesFrom*;

3. **Минимизация используемого словаря OWL.** Данная рекомендация означает описание предметной области наиболее простыми конструкциями языка. Причиной этому служит необходимость обеспечения однозначного и максимально прозрачного анализа онтологий на стыке дисциплин. Машины вывода способны обработать значительное количество сложных логических и теоретико-множественных структур, сводя их к проверке истинности или ложности десятков (или даже сотен) сгенерированных атомарных выражений (например, дизъюнктов в методе резолюции), семантика которых сложно ассоциируется с какой-либо предметной областью или классом. Для междисциплинарного обобщения же необходим более наглядный метод анализа, результатом которого должны быть понятия или правила, принадлежащие одной из исходных онтологий (а значит и дисциплин), или легко воспринимаемые пользователем системы (например, экспертом в одной из исследуемых предметных областей) как сущности на границе дисциплин. Так, целесообразно ограничивать синтаксис OWL к одному из подмножеств описательной логики [Baader, 2010]. Например, поскольку основным назначением научных онтологий предметных областей является представление понятий, но не экземпляров, возможно использование только элементов TBox [Baader, 2010], т.е. ограничиться минимальным онтографом  $\langle X, R \rangle$ .

Данный список рекомендаций не является полным, однако следования им значительно упрощают анализ онтологий предметных областей на междисциплинарном уровне, что в свою очередь, делает возможным его автоматическую реализацию, простейшим и первым необходимым вариантом которого является системная интеграция онтологий. С другой стороны, указанные рекомендации предоставляют достаточный инструментарий для создания онтологий (как в ручном, так и автоматическом режиме), используя лишь небольшое подмножество элементов OWL, что делает их совместимыми с существующими системами визуализации и редактирования онтологий [Semantic Web Wiki].

---

## Выводы

Эволюция моделей и языков представления знаний в глобальных масштабах привела к объективной необходимости исследований в области онтологического инжиниринга на междисциплинарном уровне. В результате анализа причин возникновения той или иной технологии в этом направлении установлено, что перспективным является использование языка OWL как стандартного в глобальной сети нового поколения – Semantic Web. Однако неисследованная область применения данного языка в онтолого-

---

ориентированных междисциплинарных исследованиях значительно затрудняет практические шаги в этом направлении.

Первым необходимым на данном пути этапом, проделанным в работе, является краткое описание модели и возможностей языка OWL с акцентом на его использовании на междисциплинарном уровне. Помимо этого, опыт в области системной интеграции онтологических знаний дал возможность предложить общие и практические рекомендации по созданию онтологий предметных областей, подходящих для автоматического анализа в рамках трансдисциплинарных исследований. Указанные рекомендации также могут быть полезны при разработке систем автоматического построения онтологий (например, из естественно-языкового текста), так как упрощают модель онтологии без значительного ущерба в отношении выразительной способности. Важным является и вытекающий из анализа OWL факт совместимости таких онтологий с наиболее общей моделью представления знаний в Semantic Web – моделью RDF/RDFS, обеспечивающий перспективу их использования в глобальной сети третьего поколения.

Перспективное развитие вопроса использования OWL в трансдисциплинарных исследованиях включает три основных направления:

1. Эмпирические исследования по созданию и экспертному анализу онтологий, принадлежащих разным областям знаний;
2. Адаптация существующих программных систем генерации и интеграции онтологий к междисциплинарному уровню, следуя обозначенным рекомендациям;
3. Разработка методики (возможно, не полностью формальной) создания и анализа онтологий, ориентированных на использование в трансдисциплинарных исследованиях.

---

## Литература

- [Baader, 2010] F.Baader, D.Calvanese, D.L.McGuinness, D.Nardi, P.F.Patel-Schneider. The Description Logic Handbook. Theory, Implementation and Applications. 2nd Edition. Cambridge University Press, 2010. – 624 p.
- [Connolly, 2001] D.Connolly, F.van Harmelen, I.Horrocks, D.L. McGuinness, L.A.Stein, Lucent Technologies, Inc. DAML+OIL (March 2001) Reference Description. 2001. <http://www.w3.org/TR/daml+oil-reference>
- [DARPA] The DARPA Agent Markup Language Homepage. <http://www.daml.org/index.html>
- [Duerst, 2005] M.Duerst, M.Suignard. RFC 3987: Internationalized Resource Identifiers (IRIs). IETF, January 2005. <http://www.ietf.org/rfc/rfc3987.txt>
- [Fensel, 2001] D.Fensel, F.van Harmelen, I.Horrocks, D.L.McGuinness, P.F.Patel-Schneider. OIL: An Ontology Infrastructure for the Semantic Web. // IEEE INTELLIGENT SYSTEMS, The Semantic Web 2001, 2001. – pp. 38-45
- [Genesereth] M.R.Genesereth. Knowledge Interchange Format (draft proposed American National Standard). <http://logic.stanford.edu/kif/dpans.html>
- [Hoekstra 2009. R.Hoekstra. Ontology Representation: Design Patterns and Ontologies that Make Sense. IOS Press, 2009. – 236p.
- [Lacy 2005] L.W.Lacy. OWL: Representing Information Using the Web Ontology Language. Trafford Publishing, 2005. – 205p.
- [McIlraith, 2001] S.A.McIlraith, T.Cao Son, H.Zeng. Mobilizing the Semantic Web with DAML-Enabled Web Services, 2001. <http://www.csd.abdn.ac.uk/ebiweb/papers/mcilraith.doc>
- [Motik, 2007] B.Motik. On the Properties of Metamodeling in OWL. // Journal of Logic and Computation, 17(4), 2007. – pp 617-637

- [Motik, 2012] B.Motik, B.C.Grau, I.Horrocks, Z.Wu, A.Fokoue, C.Lutz. OWL 2 Web Ontology Language Profiles (Second Edition). W3C Recommendation 11 December 2012, 2012. <http://www.w3.org/TR/2012/REC-owl2-profiles-20121211/>
- [OWL2 Overview, 2012] OWL 2 Web Ontology Language Document Overview (Second Edition). W3C Recommendation 11 December 2012, 2012. <http://www.w3.org/TR/2012/REC-owl2-overview-20121211/>
- [OWL2 Structure, 2012] OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax (Second Edition). W3C Recommendation 11 December 2012, 2012. <http://www.w3.org/TR/2012/REC-owl2-syntax-20121211/>
- [Patel-Schneider, 2012] P.F.Patel-Schneider, B.Motik. OWL 2 Web Ontology Language RDF-Based Semantics (Second Edition). W3C Recommendation 11 December 2012, 2012. <http://www.w3.org/TR/2012/REC-owl2-mapping-to-rdf-20121211/>
- [Schneider, 2003] L.Schneider. How to Build a Foundational Ontology: The Object-Centered High-level Reference Ontology OCHRE // KI 2003: Advances in Artificial Intelligence. Lecture Notes in Computer Science. Volume 2821, 2003. – pp 120-134
- [Semantic Web Wiki] Semantic Web Wiki. Tools. <http://semanticweb.org/index.php?title=Tools&oldid=53928>
- [Sowa, 2000 (1)] J.F.Sowa. Conceptual Graph Standard (updated version from 6 December 2000). NCITS.T2 Committee on Information Interchange and Interpretation. <http://users.bestweb.net/~sowa/cg/cgdpans.htm>
- [Sowa, 2000 (2)]. J.F.Sowa] Knowledge Representation: Logical, Philosophical and Computational Foundations. Brooks Cole Publishing Co., Pacific Grove, CA, 2000. – 594 p.
- [W3C Wiki] Ontology repositories. W3C Wiki. [http://www.w3.org/wiki/index.php?title=Ontology\\_repositories&oldid=56154](http://www.w3.org/wiki/index.php?title=Ontology_repositories&oldid=56154)
- [Yergeau, 2003] F.Yergeau. RFC 3629: UTF-8, a transformation format of ISO 10646. IETF, November 2003. <http://www.ietf.org/rfc/rfc3629.txt>
- [Найханова, 2005] Л.В.Найханова. Основные аспекты построения онтологий верхнего уровня и предметной области. // В сборнике научных статей "Интернет-порталы: содержание и технологии". Выпуск 3. / Редкол.: А.Н. Тихонов (пред.) и др.; ФГУ ГНИИ ИТТ "Информика". - М.: Просвещение, 2005. - С. 452-479
- [Палагин, 2010] А.Палагин, А.Михайлюк, В.Величко, Н.Петренко. К интеграции онтологий предметных областей. // Information Models of Knowledge. ITHEA, Kiev, Ukraine – Sofia, Bulgaria, 2010. – 470 с. С. 69-85
- [Палагин, 2006] О.В.Палагин, М.Г.Петренко. Модель категоріального рівня мовно-онтологічної картини світу // Математичні машини і системи. – 2006. - №3. – С. 91-104

---

## Информация об авторах

---



**Андрей Васильевич Михайлюк** – Ин-т кибернетики им. В.М. Глушкова НАН Украины, Киев-187 ГСП, 03680, просп. акад. Глушкова, 40; e-mail: [fruler@ukr.net](mailto:fruler@ukr.net)

*Основные области научных исследований: формальные модели представления знаний, логико-онтологические системы обработки знаний*

## OWL as a Standard Model for Transdisciplinary Knowledge Representation in Semantic Web

**Andrey Mihailiuk**

**Abstract:** *The reasons for emerging OWL as a standard for ontology representation in the Semantic Web are discussed in the paper. OWL structure and possibilities for representing ontologies of transdisciplinary knowledge are outlined. Recommendations for using OWL at the interdisciplinary level are proposed.*

**Keywords:** *OWL, interdisciplinary research, ontology*