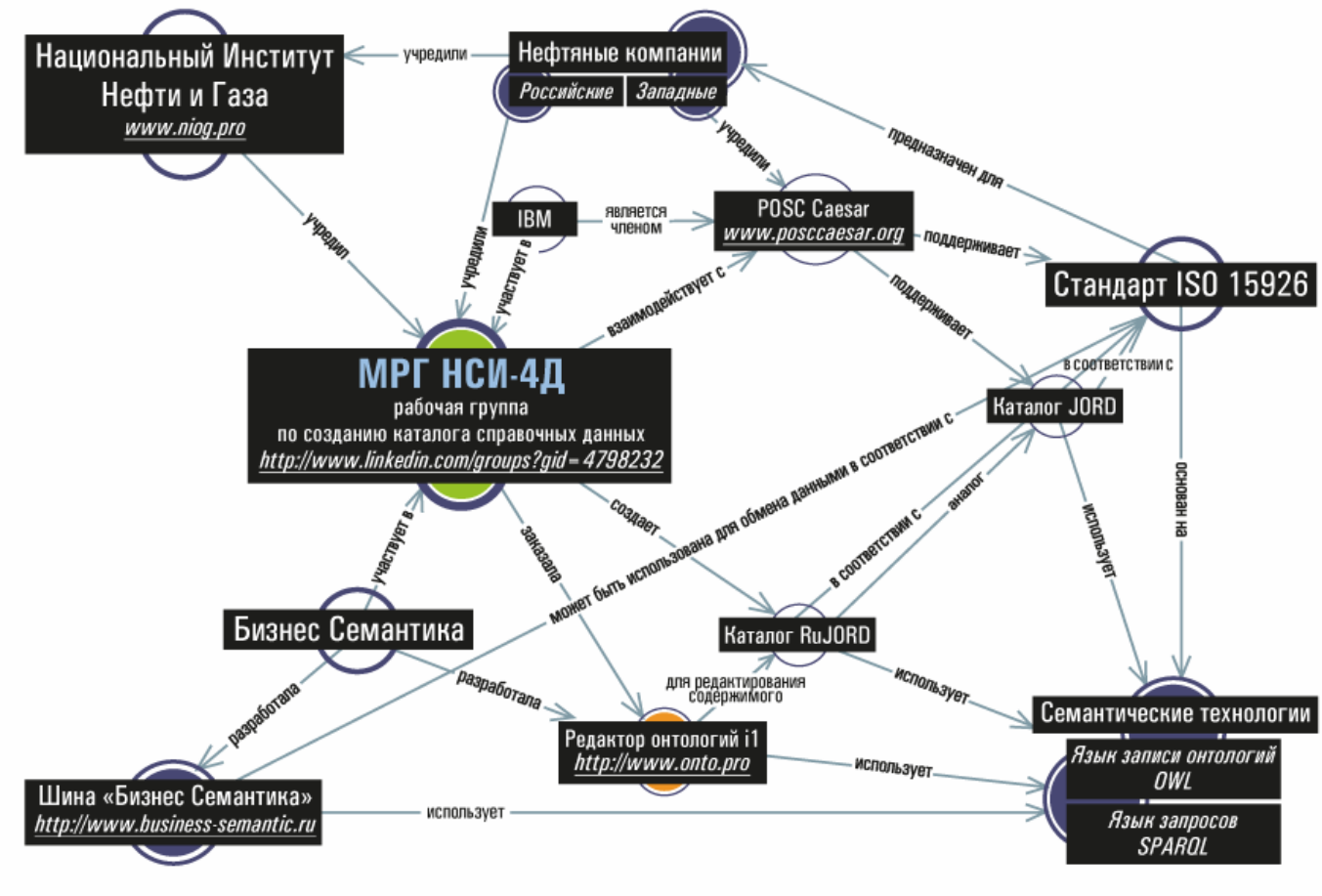
Формирование RDL в соответствии со стандартом ISO 15926 на платформе onto.pro

БИЗНЕС СЕМАНТИКА

В этом документе рассматривается наполнение данными RDL (Reference Data Library, библиотека справочных данных), при помощи программного продукта **onto.pro**.
Этот продукт создается исходя из требований проекта "Разработка отраслевой библиотеки справочных данных и отраслевого тезауруса. Модель отраслевого обмена данными".
Проект реализует Межотраслевая рабочая группа "Нормативно-справочная информация в 4D" (МРГ НСИ-4Д) при Национальном Институте Нефти и Газа.



Цели перевода НСИ в формат, соответствующий ISO 15926:

Гармонизация справочных данных с международным стандартом

Соответствие НСИ стандарту ISO 15926 позволит строить обмен информацией с любыми субъектами международного рынка.

Повышение ценности справочных данных

Данные обогащаются связями и описаниями, в результате каталог НСИ превращается в wiki-подобную семантическую базу знаний.

Расширение возможностей программного обеспечения

Перевод НСИ в семантическую форму открывает путь к разработке и внедрению программного обеспечения, основанного на семантических технологиях.
В первую очередь – это средства автоматизированного обмена информацией, сбора и интеграции данных из гетерогенных источников, аналитические приложения.

Использование RDL

Результатом ввода данных в onto.pro является онтологическая информационная модель НСИ, размещенная на программной платформе onto.pro.

Способы использования результатов:

1. Просмотр человеком в wiki-подобном представлении (база знаний):

Нагнетательная скважина

Нагнетательная скважина — Скважина, предназначенная для закачки в продуктивные горизонты воды (реже воздуха, газа) с целью поддержания пластового давления и продления фонтанного периода разработки месторождений, увеличения дебита эксплуатационных скважин, снабженных насосами и воздушными подъемниками.

Подклассы

- Скважина

Надклассы

- Водонагнетательная скважина
- Газонагнетательная скважина

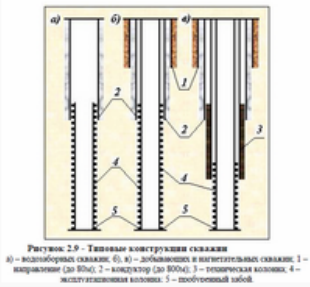
Классифицирован

- Скважина - назначение

Файлы

- *16.06.2013* Горшков - injection_well.jpg

Нагнетательная скважина



Тип:	ClassOfInanimatePhysicalObject
Дата создания:	14.06.2013
Создатель:	Gorshkov
Статус:	Draft
Ссылка на элемент JORD:	INJECTION WELL

(доступно по адресу http://www.onto.pro/RDL/RURDS000030012)

Со страницы свойств данного элемента RDL мы можем перейти на любые другие связанные с ним элементы; например, если мы находимся на странице, посвященной определенному классу труб, мы сможем перейти с нее на страницу с описанием ГОСТ, которому эти трубы соответствуют. Оттуда мы сможем скачать файл с текстом ГОСТ, увидеть другие соответствующие ему типы труб, и др. Каталог НСИ, представленный в RDL, становится многомерной структурой, содержащей множество связей между элементами, что существенно повышает его информационную ценность.

Пользователи каталога могут оставлять свои комментарии к его содержимому.

2. Использование программными системами – обращение с запросами к SPARQL-фасаду. Например, чтобы получить в виде XML результат SPARQL-запроса по всем свойствам данного вида скважин, приложение должно обратиться к URL:

http://195.189.110.39:8080/ds/query?query=SELECT+*+WHERE+{+%3Chttp%3A%2F%2Fwww.onto.pro%2FRDL%2FRURDS000030012%3E+%3Fprop+%3Fvalue+}&output=xml

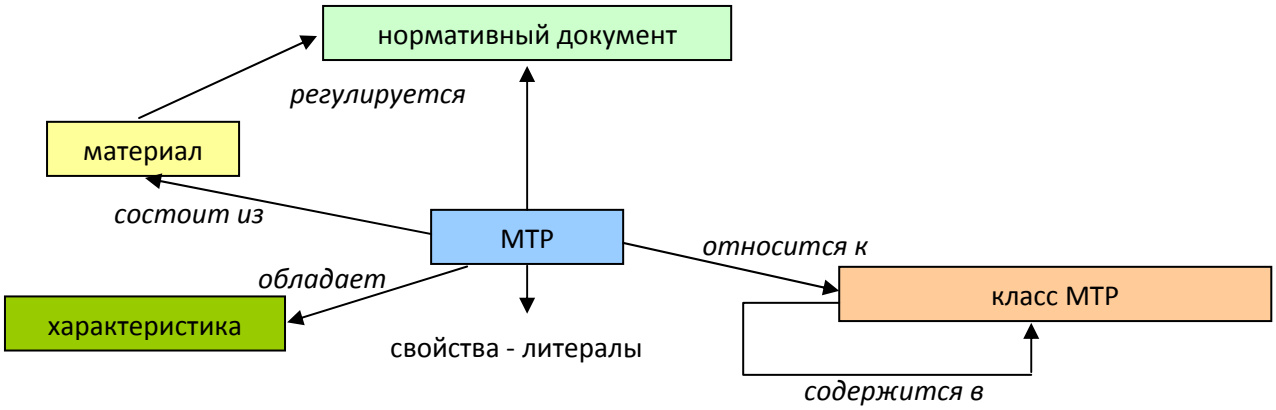
3. Использование для моделирования проектных данных (в рамках onto.pro или любого другого семантического приложения). Проектные данные могут использоваться в аналитических приложениях, основанных на семантических технологиях, применяться для обмена информацией между приложениями.

Построение информационной модели

Прежде всего, необходимо составить онтологическую модель тех данных, которые предполагается вносить в RDL. Рассмотрим пример проектирования модели для каталога МТР (материально-технических ресурсов). Пусть информацию о МТР можно представить так:

- Основные атрибуты МТР (код, название);
- Положение МТР в иерархии;
- Материал(ы), из которого состоит МТР;
- Ссылки на нормативные документы, регулирующие требования к МТР и материалу;
- Характеристики МТР.

В семантической модели каждый тип сущности должен быть представлен отдельным типом узла графа. Каждому узлу графа приписываются определенные свойства, между узлами графа устанавливаются связи. В данном случае, эскиз графа выглядит так:



Теперь эту «простую» семантическую модель нужно выразить в соответствии со стандартом ISO 15926. Стандарт подразумевает, что все узлы модели классифицируются в соответствии со стандартным набором типов, определенных в Части 2 (ClassOfIndividual). Отношения между узлами выражаются при помощи отдельных объектов – связей, или классов связей (Relationship или ClassOfRelationship). Каждая связь имеет свойства, которые при создании экземпляра связи получают значения – идентификаторы тех или иных объектов. Кроме того, связь может иметь текстовое описание.
Все имеющиеся в нашей модели элементы и связи имеют мэппинг (сопоставление) с типами ISO 15926 один-к-одному. Приведем таблицу соответствия классов нашей семантической модели типам ISO 15926.

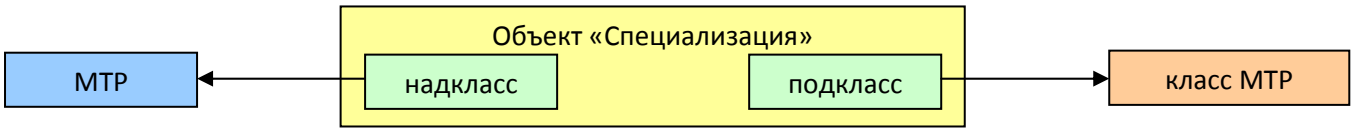
Класс модели	**Класс ISO 15926**	**Перевод**
МТР	ClassOfInanimatePhysicalObject	Класс неодушевленных предметов
Узел иерархии каталога	ClassOfInanimatePhysicalObject	Класс неодушевленных предметов
Нормативный документ	ClassOfInformationObject	Класс информационных объектов
Материал	ClassOfCompound	Класс состава
Числовая характеристика (например, диаметр)	SinglePropertyDimension	Одномерное свойство
Характеристика – перечисление (например, способ изготовления)	EnumeratedPropertySet	Свойство-перечисление

Для присвоения значений характеристикам МТР должны использоваться экземпляры шаблонов (определения шаблонов даются в Части 7 стандарта ISO 15926). В данном случае нам подходит шаблон ClassOfIndirectPropertyWithPointValue, предназначенный для отражения того факта, что характеристика X объекта Y принимает значение Z.

Осталось разобраться с тем, как выражать связи между классами нашей модели. Сведения об этом приведены в следующей таблице.

Класс 1	**Класс 2**	**Вид связи**	**Перевод**
МТР	Масса	ClassOfIndirectProperty	Класс неявного свойства
МТР	Характеристика	ClassOfIndirectProperty	Класс неявного свойства
МТР	Класс иерархии	Specialization	Специализация
МТР	Нормативный документ	Classification	Классификация
МТР	Материал	ClassOfCompositionOfIndividual	Класс состава
Материал	Нормативный документ	Classification	Классификация

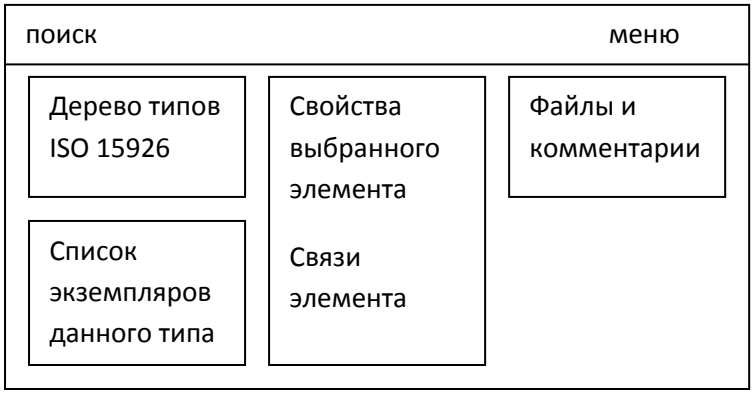
Таким образом, каждое ребро показанного выше графа превращается в отдельный объект (узел графа), имеющий, в свою очередь, две связи. Проиллюстрируем это на примере отношения Specialization, которое устанавливается между МТР и классом иерархии МТР, к которому он относится. Отношение Specialization имеет два свойства – hasSuperclass (надкласс) и hasSubclass (подкласс). Первое свойство принимает значение, равное классу иерархии МТР, второе – самому МТР. Это можно проиллюстрировать такой диаграммой:



Остальные отношения выстраиваются похожим образом. Разные виды связей имеют разные свойства. Например, связь ClassOfCompositionOfIndividual имеет свойства hasClassOfWhole и hasClassOfPart, указывающие, соответственно, на часть и целое.

Ввод информации в RDL

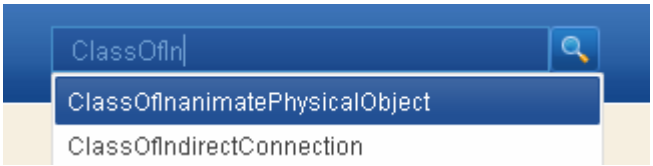
Проиллюстрируем на примерах процесс ввода данных в RDL при помощи onto.pro.
Прежде всего, опишем компоновку интерфейса onto.pro. Экран разделен на следующие части:



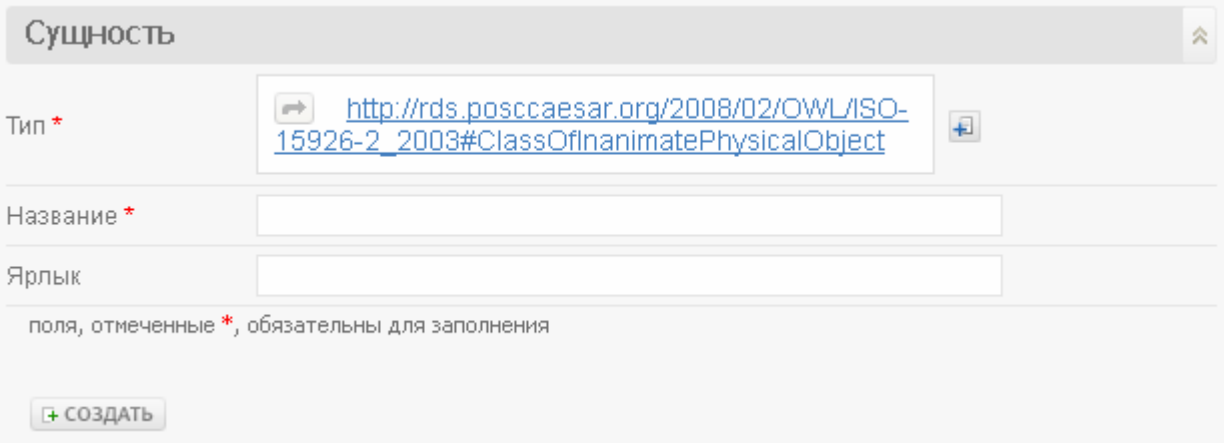
Выбранным элементом может быть как тип, так и конкретный экземпляр.

Для начала работы в onto.pro, мы должны выбрать точку доступа. Пользователи, зарегистрировавшиеся в onto.pro, получают в свое распоряжение точку доступа для экспериментов – «песочницу», sandbox. Ее и нужно выбрать.

Затем необходимо выбрать тот тип из дерева, с данными соответствующими которому мы будем работать. Наша задача – ввести данные по МТР; значит, мы будем работать с ClassOfInanimatePhysicalObject. Можно найти этот элемент в дереве, а можно воспользоваться «живым поиском», расположенным в верхней части страницы:



После выбора этого элемента, в форме в средней части страницы появятся его свойства (описание, пример использования и т.д.), а слева внизу – список уже существующих в RDL элементов этого типа. Нам надо создать новый элемент, поэтому нажмем кнопку + в списке внизу. В середине страницы появится форма создания элемента:



Тип создаваемого объекта уже выбран – ClassOfInanimatePhysicalObject (хотя, при желании, мы можем его изменить). В поле «Название» нам нужно указать уникальный идентификатор объекта. В каждом RDL должен существовать регламент присвоения таких идентификаторов, и, возможно, программный сервис для их присвоения (мы планируем создать такой сервис в onto.pro). Пока же введем идентификатор вручную, например, RURDS10010001. В поле «ярлык» надо ввести название элемента на русском языке. Создадим класс иерархии – введем значение «Скважина». После нажатия кнопки «Создать», элемент RDL будет создан.

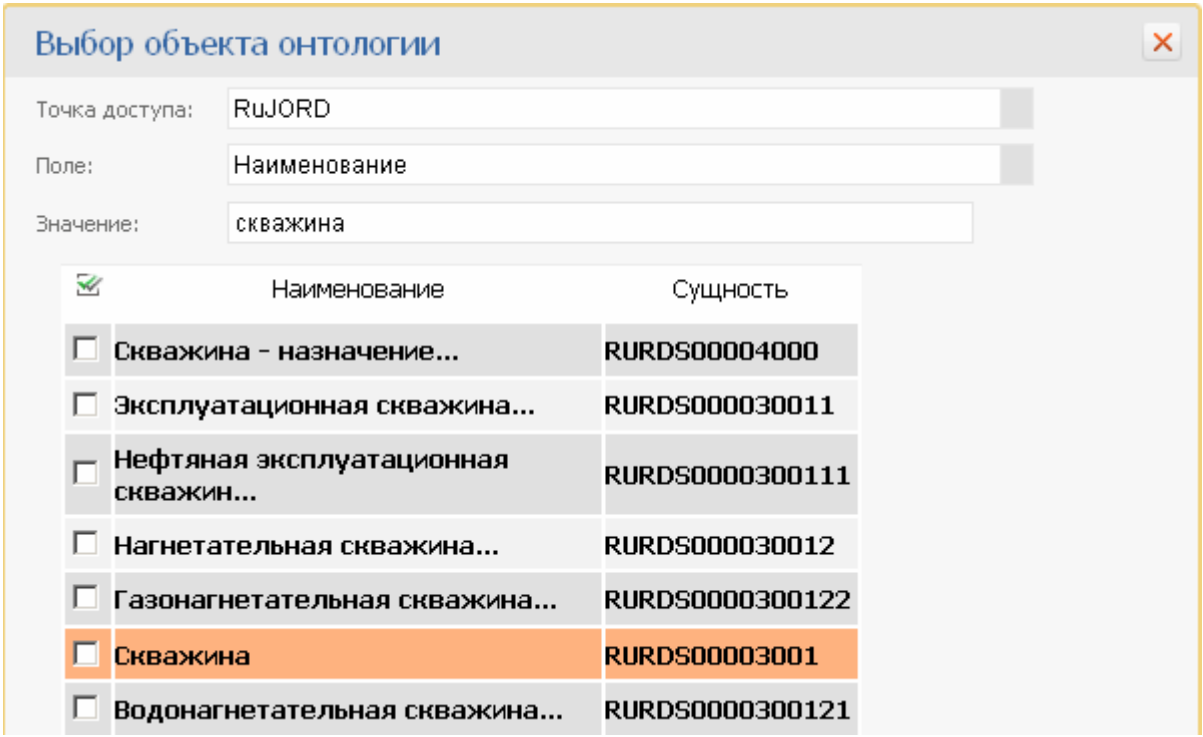
В средней части страницы появится форма редактирования созданного элемента. В ней, в области «Информация каталога JORD», перечислены свойства, которые унаследованы onto.pro из структуры JORD. Здесь можно ввести различные сведения, и сохранить форму. В правой колонке расположены формы ввода пользовательских комментариев и прикрепления файлов – то и другое отображается на странице wiki, посвященной данному элементу. Кстати, мы можем перейти к этой странице по ссылке, расположенной вверху средней части страницы.

Затем точно таким же образом создадим второй элемент типа ClassOfInanimatePhysicalObject – «Нагнетательная скважина». В списке слева внизу страницы мы видим созданные нами элементы:

☐	Нагнетательная скважина	RURDS000030012
☐	Скважина	RURDS00003001

Мы можем щелкнуть на любой из них, и увидеть его свойства в средней части страницы.

Теперь приступим к созданию связей. Выразим тот факт, что нагнетательная скважина является подвидом скважины (точно также мы могли бы связать МТР и контейнер иерархии МТР).
Для этого предназначена форма «Создать отношение» в нижней части страницы. Как мы помним, отношение с элементом иерархии выражается связью типа «Специализация». Находясь на странице свойств нагнетательной скважины, выберем тип отношения «Является подклассом», и в поле «Связанный объект» выберем «Скважина». Для этого нужно нажать на кнопку , затем в появившемся диалоговом окне найти термин «скважина», путем ввода части названия:

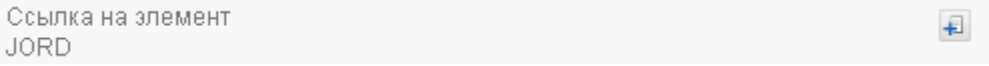


Поле «название отношения» можно оставить пустым, а можно ввести туда описание отношения – например, «подвид». Форма создания связи будет иметь следующий вид:

После нажатия кнопки «Сохранить» связь будет создана, и появится на странице:

По ссылке на уникальном идентификаторе мы можем перейти к редактированию элемента «скважина».

Далее, для созданных терминов мы должны установить мэппинг с элементом JORD. Для этого предназначена строка в форме свойств элемента:



Находясь на странице нагнетательной скважины, при помощи уже знакомой кнопки и диалога подбора выберем в точке доступа JORD элемент INJECTION WELL (поиск может происходить

долго, до половины минуты – точка доступа РСА не отличается хорошей производительностью; для решения этой проблемы, в onto.pro будет реализовано кэширование).

Чтобы проиллюстрировать создание связей других видов, добавим в каталог пример МТР - элемент «трубы нержавеющие». Его связи с нормативными документами будем устанавливать путем выбора типа связи «Классифицирован», и выбором соответствующего объекта-ГОСТа из RuJORD Характеристики (для примера возьмем диаметр) привяжем путем создания связи типа «Класс неявного свойства – класс владельца». Материал (сталь) нужно создать в справочнике, выбрав в дереве тип ClassOfCompound, а затем создать связь МТР с материалом путем выбора связи «Состав объекта – целое». После этого наша модель в целом создана; таблица связей на странице свойств трубы примет следующий вид:

Классифицирует

ГОСТ 8732-78	http://www.onto.pro/RDL/RURDS000110001

Обладатель свойства

Масса	http://www.onto.pro/RDL/RURDS00002001
Наружный диаметр	http://www.onto.pro/RDL/RURDS00002003

Целое

Сталь нержавеющая	http://www.onto.pro/RDL/RURDS10010002

Формулировки названий областей в этой форме пока условны – они нуждаются в сверке с русским переводом Части 2 стандарта ISO 15926. Тем не менее, требуемая нам информационная модель построена, и сведения о МТР выражены в соответствии с этой моделью.

Осталось рассмотреть вопрос о том, как выражать информацию о значениях характеристик МТР. Для этого, как мы уже говорили, потребуется создание экземпляров шаблона ClassOfIndirectPropertyWithPointValue. Чтобы сделать это, нам нужно перейти в onto.pro в точку доступа, соответствующую Части 7 стандарта ISO 15926, и найти в списке соответствующий шаблон. Мы увидим таблицу ролей этого шаблона:

Шаблон (ISO 15926 часть 7)

Название: **ClassOfIndirectPropertyWithPointValue**

Роли шаблона:

Название роли	Описание роли	Тип содержимого	Ограничения
hasCardinalityOfPossessorType		Cardinality	ровно 1 раз
hasUrClass		ClassOfIndividual	ровно 1 раз
hasIndirectPropertyType		ClassOfIndirectProperty	ровно 1 раз
hasDefined		ClassOfIndirectProperty	ровно 1 раз
hasScale		Scale	ровно 1 раз
hasPossessorType		ClassOfIndividual	ровно 1 раз
valPointValue		float	ровно 1 раз
hasBasePropertyType		PropertySpace	ровно 1 раз

Напротив каждого элемента (роли) шаблона находится подсказка, показывающая, какого типа значение принимает эта роль.

Внизу страницы находится форма создания экземпляра шаблона, где мы должны выбрать значения для каждой из его ролей. Этот функционал в onto.pro пока находится в стадии

разработки, и создать шаблон мы не сможем; однако, после того, как эта возможность будет реализована, мы получим возможность выразить информацию и о значениях характеристик нашего элемента справочника (а значит, и отобразить ее в wiki, а также предоставить для использования сторонним приложениям).

В данном случае, мы проделали операции по наполнению каталога вручную. Однако можно автоматизировать их при помощи шины «Бизнес Семантика», особенно если речь идет о переносе в семантическую форму уже существующего корпоративного RDL. В этом случае нужно будет создать зеркало таблиц MDM, например, в базе данных Oracle. В «Бизнес Семантике» потребуется один раз настроить соответствующие мэппинги; в дальнейшем, при изменении данных в MDM, «Бизнес Семантика» будет автоматически реплицировать данные в семантическую модель, преобразуя их в соответствии с заданными правилами. Результирующая структура данных в онтологической модели не будет отличаться от смоделированной в данном эксперименте, но на ее поддержку не потребуется никаких усилий и затрат рабочего времени.

Задать все интересующие Вас вопросы и пообщаться с Сергеем Горшковым, руководителем проекта onto.pro, вы можете:

по телефону: +7 912 23 21 108
по электронной почте: serge@business-semantic.ru

Более подробная информация о компании и сервере «Бизнес Семантика» представлена на сайте: www.business-semantic.ru
Инструмент onto.pro доступен по адресу: www.onto.pro