

# Управление семантическими активами и их повторное использование для решения задач информационного взаимодействия

© Ю. М. Акаткин    © Е. Д. Ясиновская    © М.Г. Бич    © А.В. Шилин  
Российский экономический университет им. Г.В. Плеханова  
г. Москва

[uakatkin@yandex.ru](mailto:uakatkin@yandex.ru) [elena@semanticpro.org](mailto:elena@semanticpro.org) [misha@e-projecting.ru](mailto:misha@e-projecting.ru) [a.shilin@e-projecting.ru](mailto:a.shilin@e-projecting.ru)

## Аннотация

В статье рассмотрены современные подходы к управлению семантическими активами. Предложены методы управления жизненным циклом семантических активов и их повторного использования для достижения семантической интероперабельности при взаимодействии информационных систем, моделировании процессов информационного обмена, генерации семантических веб-сервисов и переводе открытых данных в связанные открытые данные (Linked Open Data). Открытая платформа коллективной работы «Центр семантической интеграции» – проект РЭУ им. Плеханова, направленный на поддержку исследований и апробацию предлагаемых методов на базе принципов Model Driven Architecture (архитектуры, управляемой моделью, MDA).

## 1 Введение

Вопросам применения семантических методов интеграции данных из гетерогенных источников посвящено значительное число научных работ [4, 8, 9, 14, 21, 30, 31]. Эти методы более 15 лет успешно развиваются в Европе и США [1, 10] для обеспечения семантической интероперабельности информационных систем (ИС), например, в сфере электронного правительства.

Семантическая интеграция базируется на использовании семантических моделей данных – семантических активов (СА) [30] – глоссариев, словарей, таксономий, тезаурусов и онтологий. Построение СА выполняется, как правило, коллективами специалистов в ходе разработки ИС или моделирования той или иной предметной

области.

Для накопления и распространения знаний, инкапсулированных в СА, используются решения по каталогизации СА, которые обеспечивают их доступность и повторное использование. В условиях трансграничного информационного взаимодействия каталоги СА учитывают также особенности локализации. Вместе с тем существующие платформы каталогизации семантических активов не обеспечивают управление жизненным циклом (ЖЦ) СА, хотя в рамках общих рекомендаций (JOINUP [7], ADMS [23]) методология управления ЖЦ представлена.

Весьма детально управление ЖЦ метаданных регламентировано стандартами ISO/IEC 11179 [6] и соответствующим ГОСТ Р ИСО/МЭК 11179-1-2010 «Информационная технология. Регистры метаданных (РМД)» [18]. Однако они не учитывают особенностей управления семантическими активами.

Поскольку СА являются не только семантическими моделями, но и стандартами данных, целесообразно обратить внимание также на рекомендации W3C [15] по управлению ЖЦ разработки и распространения стандартов в сфере веб-технологий.

Таким образом, для решения задачи управления ЖЦ СА необходимо развитие существующих методологических разработок и действующих стандартов для учета специфики процессов разработки СА, их локализации и повторного использования.

Заметим, что потребность в повторном использовании ранее разработанных СА возрастает по мере интеграции все большего числа разнородных ИС, поскольку позволяет избежать крайне затратных и дублирующих исследований. При переводе открытых данных в связанные открытые данные, работа экспертов направлена на применение ранее разработанных СА для придания данным смысловой окраски.

В то же время при моделировании процессов информационного обмена уже разработанных СА,

как правило, не хватает и приходится решать задачу построения более детальной семантической модели обмена применительно к предметной области взаимодействующих ИС. Вместе с тем, в процессе формализации предметной области зачастую возникает «разрыв» при взаимодействии экспертов домена и разработчиков ИС, что приводит к множественным доработкам и получению неудовлетворительных результатов. Сочетание методов коллективной работы и принципов MDA – архитектуры, управляемой моделью, предложенной OMG (Object Management Group) [11] – призвано сократить этот разрыв.

Проект РЭУ им. Г.В. Плеханова по созданию Центра семантической интеграции (ЦСИ) включает макетирование открытой платформы коллективной работы на базе существующего международного и российского опыта, разработки и апробации методов управления семантическими активами, реализации технологий повторного использования СА для решения прикладных задач информационного взаимодействия.

## 2 Управление семантическими активами

В рамках европейской программы интероперабельности ISA [1] для каталогизации СА разработан стандарт ADMS (Asset Description Metadata Schema) [23], который является профилем Словаря каталогов данных DCAT (Data Catalogue Vocabulary) [29]. Причем если DCAT предназначен для упрощения взаимодействия между каталогами данных, т.е. сам каталог находится в непосредственно центре словаря, то ADMS ориентирован на СА в каталоге.

ЖЦ СА в ADMS и, соответственно, в платформе JOINUP описывается 4-мя статусам: готов, устарел, в разработке, изъят. То есть управление ЖЦ на стадии разработки СА остается за рамками платформы JOINUP.

В ЦСИ стандарт ADMS также применяется для каталогизации СА, хранения и публикации их описаний. В то же время ЖЦ СА должен учитывать стадии коллективной работы над СА, проведения экспертизы и оценки качества опубликованного семантического актива. Поэтому в рамках нашего исследования были рассмотрены как рекомендации W3C [15], так и стандарты серии ГОСТ Р ИСО/МЭК 11179 [18], включая не вошедший в российскую версию ISO/IEC 11179-6:2015 [6].

При разработке и распространении стандартов W3C используются следующие стадии ЖЦ:

**Рабочий проект (WD)** - документ, который W3C опубликовал для рассмотрения сообществом, в том числе членами W3C, сообществами и другими техническими организациями. Рабочие проекты не обязательно представляют собой консенсус Рабочей группы, и не подразумевают какого-либо одобрения W3C.

**Кандидат на рекомендацию (CR)** – документ, который удовлетворяет техническим требованиям Рабочей группы и прошел стадию экспертного обсуждения в сообществе.

**Предлагаемая рекомендация (PR)** – документ, который по качеству соответствует требованиям W3C и был принят директором W3C.

**Рекомендация W3C (REC)** – спецификация, набор руководящих принципов или требований, которые, после проведения обсуждения получили одобрение членов W3C и директора W3C.

Следует подчеркнуть, что стандарты в сфере веб-технологий рассматриваются в W3C как единый документ. В ЦСИ такой подход может быть принят только для управления описаниями семантических активов. Однако для решения задач применения и повторного использования семантических активов, которым посвящен раздел 3 данной статьи, только описаний («карточек») СА недостаточно – требуется публикация и распространение содержания семантического актива, его элементов.

ISO/IEC 11179-6:2015 регулирует процесс регистрации регистров метаданных и входящих в них элементов (записей регистра) в соответствии со следующими стадиями ЖЦ:

**Предпочтительный стандарт (Preferred Standard)**. Регистрирующий орган подтвердил, что запись в регистре предпочтительна для использования в сообществе пользователей регистра метаданных.

**Стандарт (Standard)**. Регистрирующий орган подтвердил, что запись в регистре необходимого качества и имеет широкий интерес для использования в сообществе пользователей регистра метаданных.

**Надлежащего качества (Qualified)**. Регистрирующий орган подтвердил, что обязательные атрибуты метаданных являются полными, обязательные атрибуты метаданных соответствуют применяемым требованиям к качеству.

**Зарегистрирован (Recorded)**. Регистрирующий орган подтвердил, что все обязательные атрибуты метаданных заполнены.

**Кандидат (Candidate)**. Запись в регистре предложена для прохождения регистрации в регистре.

**Не завершен (Incomplete)**. Отправитель хочет информировать сообщество, которое использует регистр метаданных, о существовании новой записи в регистре.

**Изъят (Retired)**. Регистрирующий орган одобрил для записи в регистре, что ее больше не рекомендуется использовать в сообществе пользователей регистра метаданных или, что она больше не должна быть использована.

**Замена (Superseded)**. Регистрирующий орган одобрил для записи в регистре, что ее больше не

рекомендуется использовать в сообществе пользователей регистра метаданных, и что теперь предпочтительно использовать заменяющую запись в регистре.

Применение этого стандарта в ЦСИ требует организации регистрирующего органа, поэтому должно быть отнесено на более поздние этапы развития Центра.

В ЦСИ представлены не только описания семантических активов, собранные в каталог, но и собственно содержание СА – набор элементов и их свойств. Поэтому в ЦСИ предлагается реализовать как ЖЦ описания («карточки») СА, т.е. записи в ADMS каталоге, так и ЖЦ содержания СА. Необходимость такого деления состоит в том, что описание актива – рассматривает СА как целую, неделимую единицу, а содержание СА может разбиваться на части (ветки, разделы, наборы, элементы), для каждой из которых может выполняться отдельный рабочий процесс.

Для описания СА может быть использован имеющий широкое распространение и признание в веб-сообщества ЖЦ W3C, поскольку он учитывает такие особенности ЦСИ, как привлечение широкого круга специалистов, разнообразие предметных областей СА, использование СА в Semantic Web и обмене информацией с использованием web-технологий. С организационной точки зрения важно отметить, что роль «Директора ЦСИ» не требуется, решение должно приниматься экспертным сообществом с использованием механизмов голосования, оценок, рейтингов и др. Семантические активы могут быть классифицированы по источникам управления. Внешние СА (например, EuroVoc) разрабатываются и развиваются за пределами ЦСИ, внутренние СА – непосредственно в ЦСИ. Внешние активы представляются в ЦСИ своими описаниями, однако в рамках ЖЦ внешних СА (например, в процессе обсуждения или экспертной оценки) могут возникнуть рекомендации по их загрузке для перевода и повторного использования. Таким образом, инициируется появление внутреннего СА, например, русской ветки EuroVoc. Внутренние СА могут создаваться участниками ЦСИ (методами ручного ввода или автоматизированной загрузки из внешних программ), а также возникают при повторном применении СА – в процессе моделирования процессов информационного обмена и при переводе открытых данных в связанные.

В ходе работы над содержанием СА (наполнение, перевод) могут возникать этапы ЖЦ, отражающие стадии коллективной работы участников рабочей группы. Например: выбор элементов СА (стадия «подготовка»), назначение заданий (стадия «перевод»), контроль исполнения заданий (стадия «проверка») и оценка качества (стадия «принят»).

Для этапов ЖЦ описания СА в соответствии со схемой [23, 28] используется свойство «adms:status»

класса «adms:Asset». А для этапов ЖЦ содержания СА – свойство «adms:status» класса «adms:AssetDistribution»

Важным этапом является экспертная оценка, уровень которой выбирается в зависимости от интереса к СА: (1) валидация в рамках рабочей группы, (2) публичная экспертиза участников сообщества, (3) независимая экспертиза, проводимая экспертами предметной области.

Переходы между этапами ЖЦ описания актива доступны только после окончания выполнения этапов ЖЦ содержания актива (рис.1). Таким образом объединяются ЖЦ описаний СА и ЖЦ процессов, выполняемых над содержимым СА. Например, невозможно будет получить состояние «Рекомендация», если ещё выполняется процесс «Экспертная оценка». При нахождении описания актива на том или ином этапе ЖЦ для изменения содержания актива могут запускаться только процессы доступные на этом этапе. Например, «Перевод» содержания актива может выполняться только, если описание актива находится на этапе «Рабочий проект».

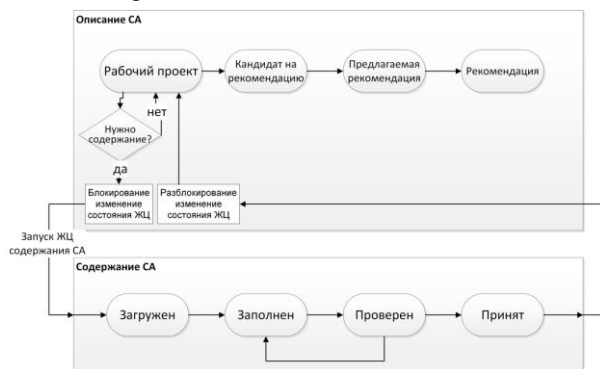


Рисунок 1 - ЖЦ описания и содержания семантического актива

С учетом особенностей семантических активов и процессов коллективной работы над их содержанием, управление СА должно происходить в рамках ЖЦ описания СА, представленного в таблице 1.

Таблица 1 - Этапы ЖЦ СА

Этап ЖЦ описания СА	Процессы работы над содержанием СА	Комментарий
<b>Рабочий проект (РП)</b>	Публикация описаний внешнего СА. Загрузка внешних СА. Создание внутреннего СА Перевод. Внесение изменений. Экспертная оценка.	Основное наполнение содержания актива (загрузка внешнего СА, создание нового СА, перевод и т.д.) и уточнение описания СА (наполнение полей, классификация,

	Удаление/снятие с публикации.	связи с другими активами).
<b>Кандидат на рекомендацию (КР)</b>	Внесение изменений Экспертная оценка Удаление/снятие с публикации	Основная экспертная оценки и внесение косметических изменений
<b>Предлагаемая рекомендация (ПР)</b>	Экспертная оценка Удаление/снятие с публикации	Решение, что СА может быть рекомендован к использованию в ЦСИ.
<b>Рекомендация (РЕК)</b>	Удаление/снятие с публикации Создание новой версии	Могут приниматься решения о создании новой версии или удалении актива из каталога.

Решение задач повторного использования СА и их элементов накладывает ограничение на удаление/снятие с публикации. Кроме очевидных механизмов уведомления участников, использующих СА об изменениях, должны быть поддержаны и методы «вывода из использования». Исследование показало, что существующие каталоги СА также обеспечивают именно их накопление, то есть параллельно могут существовать несколько версий СА. Это означает, что при интеграции гетерогенных систем для развития семантического ядра обмена должно быть обеспечено сохранение знаний об изменениях семантики, установление связей между «старыми» и «новыми» СА.

Например, СА в состоянии «Рекомендация» может быть снят с публикации или определен как устаревший. При этом он остается «Рекомендацией», т.е. состояние «Рекомендация» является финальным для текущей версии актива и не может быть изменено. Используя свойство [23, 28] «owl:versionInfo» класса «adms:Asset» можно установить флаг со значением «устаревший» и связь с указанием более свежей версии (свойство «adms:last» класса «adms:Asset»). Это позволяет обеспечить использование ранее разработанных СА и реализовать явные правила преобразований СА из устаревших в актуальные версии.

### 3 Применение семантических активов

(Повторное) использование семантических активов является базовым принципом, обеспечивающим достижение семантической интероперабельности. В проекте ЦСИ ведется исследование, апробация и развитие методов разработки и (повторного) использования СА для перевода открытых данных в связанные открытые данные, моделирования схем данных и

информационного обмена при интеграции ИС, на основе принципов коллективной работы экспертов, повторного использования СА и MDA.

#### 3.1 Перевод открытых данных в связанные

Концепция открытых данных (Open Data, OD) декларирует доступ к информации не просто в визуальных, но и в машиночитаемых форматах [17], что обеспечивает возможность ее многократного использования.

За годы продвижения принципа открытости сформирован значительный пул OD, однако практика показывает, что такой способ публикации не обеспечивает возможности использования данных в полной мере. Более того, анализ и визуализация открытых данных зачастую может выполняться только после дополнительной обработки или «ручным способом». Это связано с тем, что требования ни к структурированию и гармонизации данных, ни к описывающим их метаданным, в том числе связи с контекстом и терминами предметной области, в OD не предъявляются.

Большой интерес для мирового сообщества представляет концепция связанных открытых данных (Linked open data, LOD), которая использует [2] срез технологий Semantic Web для обеспечения совместимости, облегчения многократного использования и комбинирования данных. Семантическое связывание данных позволяет объединять данные из различных источников, не прибегая к созданию специализированных программ, и предоставлять доступ к этим данным.

LOD обеспечивают децентрализацию данных, помогают в работе с большими объемами данных [2, 3]. Способы применения LOD для решения задач интеграции данных в предметных областях с интенсивным использованием данных (Data Intensive Domains) неоднократно рассматривались коллективом авторов под руководством В.А. Серебрякова, например, [21].

Исследуемые в рамках ЦСИ методы перевода открытых данных в LOD основываются на четырех сформулированных Тимом Бернесом-Ли принципах Semantic Web [2] и на дополнительном **пятом принципе** – повторного использования семантических ресурсов. В рамках проекта ЦСИ ведется исследовательская работа по апробации и развитию методов перевода открытых данных в связанные открытые данные (OD2LOD) путем формирования метаданных на основе использования СА.

В процессе перевода OD2LOD должны быть реализованы следующие этапы: (1) подготовить данные (загрузить, нормализовать, «связать»), (2) провести их верификацию и (3) получить сформированный в рамках стандарта пакет LOD, готовых для (4) публикации и (5) последующего использования в процессах информационного обмена.

Для апробации выбранных методов разработаны примеры, реализующие шаги бизнес-процесса OD2LOD. В примерах использованы открытые данные, опубликованные на официальных порталах государственных открытых данных ([data.gov.ru](http://data.gov.ru), [mos.ru](http://mos.ru)).

Для подготовки данных в части загрузки и нормализации используется инструмент Open Refine [26], а для связывания с семантическими ресурсами апробируется платформа Silk [27]. Для хранения моделей и данных применяется Apache Jena [24]. На этапе публикации особенно важным является применение средств визуализации LOD, исследование которых сегодня активно поддерживается международным научным сообществом [5].

При отсутствии готовых решений и/или при необходимости их интеграции ведется разработка собственных приложений ЦСИ для повторного использования СА для OD2LOD и выполнения всех этапов перевода.

Таким образом, в ЦСИ будут реализованы:

- Единая точка доступа – все необходимые инструменты собраны в одном месте, имеют общий интерфейс.
- Инструменты связывания – позволяют снизить стоимость самой трудоемкой операции по подготовке данных.
- Инструменты доступа к данным – позволяют получать данные из различных источников.
- Инструменты визуализации – позволяют показать данные в различных представлениях.
- Повторное использование моделей – возможность использовать или доработать имеющиеся модели под свои нужды вместо разработки новых.
- Хранение данных – возможность автоматически настраивать хранилище и хранить большие объемы данных, снижая риски утери первоначального источника.

### 3.2 Моделирование схем информационного обмена

Гетерогенная среда информационных систем в областях с интенсивным использованием данных, которые представляют ценность для других потребителей, диктует необходимость организации взаимодействия и моделирования схем информационного обмена.

Интеграция данных на семантическом уровне поддерживает унифицированное представление данных на основе семантических свойств в контексте единой онтологии предметной области [22]. Для достижения этого должны использоваться, в том числе повторно, семантические активы. Преимущество семантического подхода к

интеграции ИС перед традиционным убедительно показано в [30].

Теоретические исследования описывают методы семантической интеграции [21, 22, 30, 31], однако на практике их применение затруднено. Несмотря на то, что «семантическое моделирование стало предметом исследований начиная с конца 1970-х годов» [4], проблема гетерогенных семантик данных, которые возникают как внутри систем, так и при обмене между ними, остается актуальной [8].

На 7-й международной конференции GRID 2016 в Дубне авторы статьи представили результаты<sup>1</sup> исследования существующих подходов к семантической интеграции и сформулировали ряд проблем, на решение которых направлен проект ЦСИ. В первую очередь это проблема организации взаимодействия между специалистами предметной области (домена) и ИТ-специалистами: моделирование информационной системы и процессов взаимодействия в гетерогенной среде должно опираться на знания о предметной области, ее объекты и связи между ними, то есть должно быть доступно и понятно всем участникам вне зависимости от сферы их компетенции.

Необходимость разработки специальных модулей импорта/экспорта данных возникает при интеграции ИС, поскольку схемы данных различных ИС обычно не соответствуют друг другу, причем источники данных и их получатели используют различные контексты. Неявные допущения, существующие в каждом источнике, должны быть явно описаны и использованы, чтобы исключить конфликты при сопоставлении и совместном использовании данных из этих систем [9].

Современные подходы к интеграции данных в основном ориентированы на использование XML (eXtensible Markup Language) форматов обмена данными, для описания которых используются XSD схемы (XML Schema definition). Наиболее популярные на сегодняшний день подходы к построению XSD схем используют модели на логическом уровне интеграции.

При этом построение моделей, как правило, проводится ИТ-специалистом, не имеющим достаточных знаний в предметной области. Моделирование происходит на конечных этапах разработки ИС в рамках платформу-зависимой реализации. Итоговый результат заказчик/потребитель может увидеть только после окончания разработки. Соответственно, такие модели семантических связей не содержат и машинное «понимание» остается невозможным.

Применение семантических технологий для реализации задач информационного обмена, в том числе семантической шины [32], требует участия в моделировании специалистов домена. Однако

<sup>1</sup> <https://indico-new.jinr.ru/indico/contributionDisplay.py?sessionId=16&contribId=62&confId=85>

ориентация инструментов разработки на ИТ-специалистов определяет высокий порог вхождения экспертов домена в процессы разработки схем обмена данными (и ИТ-систем в целом).

Одной из основных задач ЦСИ является организация совместной коллективной работы экспертов домена и ИТ-специалистов. Последовательный переход от бизнес-требований к технической реализации в ЦСИ базируется на принципах архитектуры управляемой моделью (Model Driven Architecture, MDA) [11] и (повторном) использовании СА. MDA дает возможность сохранить инвестиции, сделанные в разработку бизнес-логики даже при смене технологических платформ.

Применение этого подхода в ЦСИ позволяет использовать СА при реализации схем информационного обмена и обогащении существующих данных семантическим описанием. Примером такого обогащения является установка связи каждого элемента XSD-схемы (в частности, WSDL описании web-сервиса), который является семантическим свойством данных в предметной области, с СА или его элементами, представленными в ЦСИ. В описании XSD схемы эта информация сохраняется в элементе схемы «*xsd:documentation*».

Основой разработки СА и их применения для решения прикладных задач является реализация 4-х уровневой иерархии метамodelей (M3-M0) [12, 13, 20]. Она обеспечивает целостность моделей, создаваемых и предоставляемых в ЦСИ, следующим образом:

- в ядре платформы ЦСИ реализуется фабрика моделей (т.н. мета-объектное средство, Meta-Object Facility, MOF) – уровень M3;
- в состав ЦСИ включается метамodelь UML – уровень M2;
- для представления модели СА выбрана нотация UML (Unified Modeling Language) – уровень M1;
- СА являются экземплярами классов из моделей СА – уровень M0.

В процессе разработки СА, используя уже существующие в ЦСИ элементы СА, «предметник» должен иметь возможность сформировать структуру СА в визуальном режиме. В результате, автоматически или с минимальными уточнениями, создается общая платформо-независимая модель (Platform Independent Model, PIM) [11]. А также обеспечивается её связь с вычислительно-независимой моделью (Computation Independent Model, CIM).

Готовая PIM дополняется контекстными компонентами (например, связанными СА), формируя на выходе контекстную PIM, которая затем трансформируется в платформо-зависимую модель (Platform Specific Model, PSM)

реализуемой платформы. При необходимости замены платформы может быть проведена повторная генерация PSM на основе уже имеющейся PIM.

CIM, PIM и PSM модели могут быть визуализированы в нотации UML или разработаны во внешнем UML-редакторе, а затем загружены в ЦСИ в XMI формате. На основе PSM генерируются платформенные компоненты (создаются базы данных, схемы XSD и т.п.). Схемы XSD, обогащённые связью с СА, используются в реализации семантических сервисов (Semantic Web Services).

Следует отметить, что такой подход позволяет в перспективе включить в ЦСИ и другие метамodelи, например, CWM (Common Warehouse Metamodel). А загруженные в ЦСИ модели предметных областей, справочные модели и базовые словари, такие как ядро национальной модели обмена информацией США NIEM (National Information Exchange Model) [10], базовые европейские словари (Core Vocabularies) [25], российские справочники и классификаторы, могут быть использованы для построения схем информационного обмена и обеспечения семантической интероперабельности ИС.

## 4 Заключение

В статье рассмотрены подходы к управлению жизненным циклом семантических активов и возможности применения СА. Определена необходимость объединения ЖЦ описания СА и ЖЦ содержания СА, с учетом их особенностей, в том числе повторного использования. Для ЖЦ описания СА предложено применение методологии W3C, расширенной процессами работы над содержанием СА.

Приведены предложения по реализации перевода открытых данных в связанные открытые данные с использованием СА. Сформулированы основные функции процессов OD2LOD для их дальнейшей реализации в ЦСИ.

Рассмотрен подход к устранению «разрыва» при взаимодействии экспертов домена и разработчиков ИС в процессе семантического моделирования предметной области на основе сочетания методов коллективной работы, повторного использования СА и принципов MDA.

Проект РЭУ им. Г.В. Плеханова по созданию Центра семантической интеграции представлен как площадка для апробации методов и инструментария, разрабатываемых в ходе проведения научно-исследовательской работы. Создаваемый макет открытой платформы коллективной работы ЦСИ будет обеспечивать каталогизацию СА, управление их жизненным циклом, а также практическое применение СА для решения задач информационного взаимодействия.

При проведении прикладных исследований с использованием ЦСИ научные коллективы с привлечением молодых ученых и студентов получают возможность семантического моделирования различных предметных областей, создания глоссариев, таксономий, тезаурусов и онтологий для применения в исследовательских проектах, систематизации научных материалов и машинной обработке научных текстов на русском языке.

ЦСИ как семантическая база знаний, обеспечивающая возможность многократного повторного использования СА, может эффективно использоваться в областях с интенсивным использованием данных, таких как:

- Электронное правительство. Обеспечение коллективной работы разноплановых специалистов по созданию семантических моделей государственных услуг и организации межведомственного взаимодействия для обеспечения семантической интероперабельности в российском ЭП;
- Анализ данных. Новый уровень аналитики по связанным данным из разных источников, визуализация и сопоставление связанных показателей.
- Семантическая интеграция информационных систем. Владельцы ИС и коллективы разработчиков получают возможность моделирования предметной области ИС, построения схем информационного обмена с использованием федеративных моделей данных.

## Литература

- [1] Annex 2 to the Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of Regions 'Towards interoperability for European public services' EUROPEAN COMMISSION Bruxelles, le 16.12.2010 COM(2010) 744 final [http://ec.europa.eu/isa/documents/isa\\_annex\\_ii\\_eif\\_en.pdf](http://ec.europa.eu/isa/documents/isa_annex_ii_eif_en.pdf)
- [2] T. Berners-Lee. Linked Data – Design Issues <http://www.w3.org/DesignIssues/LinkedData.html>
- [3] Chris Bizer, Tom Heath, Tim Berners-Lee: Linked Data: Principles and State of the Art, Arpil 2008, <http://www.w3.org/2008/Talks/WWW2008-W3CTrack-LOD.pdf>
- [4] C. J. Date, An Introduction to Database Systems (8th Edition). Pearson Education Inc., 2004, p. 1024, ISBN 0-321-18956-6
- [5] Jiří Helmich, Jakub Klimek, and Martin Nečaský, Visualizing RDF Data Cube using the Linked Data Visualization Model, in The Semantic Web: ESWC 2014 Satellite Events, Anissaras, Crete, Greece, Springer Verlag, ISBN: 978-3-319-11955-7, ISSN: 0302-9743, pp. 368-373, 2014
- [6] International Standard ISO/IEC 11179-6:2015, Information technology — Metadata registries (MDR) — Part 6: Registration
- [7] Joinup, European Commission, [https://joinup.ec.europa.eu/asset/page/practice\\_aid\\_s/what-semantic-interoperability](https://joinup.ec.europa.eu/asset/page/practice_aid_s/what-semantic-interoperability)
- [8] Madnick S., Gannon T., Zhu, H., Siegel M., Moulton A., Sabbouh M., Framework for the Analysis of the Adaptability, Extensibility, and Scalability of Semantic Information Integration and the Context Mediation Approach, Massachusetts Institute of Technology Cambridge, MA, USA, 2009
- [9] Madnick, S.E., & Zhu, H., Improving data quality through effective use of data semantics, Data and Knowledge Engineering, 59(2), 2006, p.460-475
- [10] National Information Exchange Model, <https://www.niem.gov>
- [11] Object Management Group, MDA Specification, <http://www.omg.org/mda/specs.htm>
- [12] Object Management Group, Meta Object Facility (MOF) Specification, 2000, <http://www.omg.org/spec/MOF/2.5/>
- [13] Iman Poernomo, The Meta-Object Facility Typed, SAC'06 April 23-27, 2006, Dijon, France, <http://calcium.dcs.kcl.ac.uk/1259/1/acm-paper.pdf>
- [14] Walaa S. Ismail, Mona M. Nasr, Torky I. Sultan, Ayman E. Khedr, Semantic Conflicts Reconciliation as a Viable Solution for Semantic Heterogeneity Problems, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 4, No.4, 2013
- [15] W3C Consortium Process Document - <http://www.w3.org/Consortium/Process/>
- [16] Акаткин, Ю.М., Дрожжинов В.И., Ясиновская Е.Д., Что такое система межведомственного взаимодействия НИЕМ, 2014, [http://www.cnews.ru/reviews/new/ikt\\_v\\_gossektor\\_e\\_2014/articles/chto\\_takoe\\_sistema\\_mezhvedomstvennogo\\_vzaimodejstviya\\_niem/](http://www.cnews.ru/reviews/new/ikt_v_gossektor_e_2014/articles/chto_takoe_sistema_mezhvedomstvennogo_vzaimodejstviya_niem/)
- [17] Бегтин И., Дубова Н., О данных в открытую, Открытые системы № 03, 2015, <http://www.osp.ru/news/articles/2015/09/13045074/>
- [18] ГОСТ Р ИСО/МЭК 11179-1-2010, Информационная технология. Регистры метаданных (РМД)
- [19] Кузнецов М.Б., Трансформация UML-моделей и ее применение в технологии MDA, 2016, [http://citforum.ru/SE/project/uml\\_mda/](http://citforum.ru/SE/project/uml_mda/)
- [20] Марков Е., Архитектура, управляемая моделью, 2005, <http://citforum.ru/gazeta/13/>
- [21] Малахов Д.А., Серебряков В.А., Теймуразов К.Б., Шорин О.Н., Интеграция библиографических данных в Linked Open Data, Труды 16-й Всероссийской научной конференции «Электронные библиотеки:

- перспективные методы и технологии, электронные коллекции» — RCDL-2014, Дубна, Россия, 13–16 октября 2014 г.
- [22] Серебряков В.А., Семантическая интеграция данных, 2012, <http://sp.cmc.msu.ru/proseminar/2012/serebryakov.2012.04.20.pdf>
- [23] Спецификация и описание Asset Description Metadata Schema, 2016 <https://joinup.ec.europa.eu/asset/adms/home>
- [24] Спецификация и описание Apache Jena, <http://jena.apache.org/>
- [25] Спецификация и описание EU Core Vocabularies, 2015, <https://joinup.ec.europa.eu/node/145983>
- [26] Спецификация и описание Open Refine, <http://openrefine.org/>
- [27] Спецификация и описание Silk Framework, <http://silkframework.org/>
- [28] Спецификация и описание W3C Asset Description Metadata Schema (ADMS), август 2013, <http://www.w3.org/TR/vocab-adms/>
- [29] Спецификация и описание W3C Data Catalogue Vocabulary, 2014, <http://www.w3.org/TR/2014/REC-vocab-dcat-20140116/>
- [30] Черняк Л., Интеграция данных: Семантика и синтаксис, Открытые системы №10, 2009, <http://www.osp.ru/os/2009/10/11170978/>
- [31] Шибанов С. В., Яровая М. В., Шашков Б. Д., Кочегаров И. И., Трусов В. А., Гришко А. К. Обзор современных методов интеграции данных в информационных системах // НиКа. 2010. №. URL: <http://cyberleninka.ru/article/n/obzor-sovremennyh-metodov-integratsii-dannyh-v-informatsionnyh-sistemah>
- [32] Ульянов Д., Проект семантической шины semap, апрель 2007, <https://dulanov.wordpress.com/2007/04/17/proekt-semanticheskogo-menadgera-semap/> Т. Berners-Lee. Linked Data – Design Issues <http://www.w3.org/DesignIssues/LinkedData.html>

### **Management and (re)use of semantic assets for information sharing**

Yury Akatkin, Elena Yasinovskaya, Mikhail Bich, Andrey Shilin

The article represents current research results gained at The Center of Semantic Integration (Plekhanov Russian University of Economics). It shows the approach to the management of semantic assets and their (re) use for information sharing, data modelling based on Model Driven Architecture, semantic web services generation and transformation of open data to Linked open data.