

Отчет за 2017 год
по проекту 17-07-00228

МЕтодология и ТЕхнология формирования Онтологий на основе интегРации с гетерогенными источниками данных (МЕТЕОР)

Коды классификатора: **07-966, 07-971**

Ф.И.О. руководителя проекта: Гаврилова Татьяна Альбертовна

Аннотация отчета

Проект МЕТЕОР ориентирован на решение актуальной проблемы проектирования и формирования баз знаний предметных областей на основе онтологий через интеграцию и обработку структурированных массивов данных из гетерогенных источников.

Проект состоит из трех этапов:

1. Анализ состояния современных исследований в области формирования и поддержки баз знаний онтологического типа и разработка основ методологии интеграции структурированных данных из гетерогенных источников в базы знаний различных предметных областей — корпоративных, научных, учебных и др.
2. Создание МЕТЕОР-технологии формирования онтологий и интеграции структурированных данных из гетерогенных источников с учетом человеческого фактора, в частности, когнитивных аспектов дизайна и генерации онтологий.
3. Разработка архитектуры прототипной системы МЕТЕОР-демо, поддерживающей/реализующей эту технологию и апробация разработанных методологии и технологии на примере создания базы знаний учебного процесса ВУЗа.

Первый этап 2017 года был посвящен детальному анализу проблематики и изучению мирового опыта в области формирования корпоративных онтологических баз знаний и интеграции данных из гетерогенных источников. Эти источники включают реляционные базы данных, электронные таблицы, XML-файлы и т.п. Были разработаны подходы к новой методологии формирования корпоративных онтологий.

В течение этапа был проведен анализ существующих подходов к формированию корпоративных онтологических БЗ и к интеграции имеющихся данных в созданную онтологию. Были решены следующие задачи:

Раздел 1. Анализ состояния мировых научных трендов в области баз знаний онтологического типа, включающего следующие части:

- Задача 1-1.** обзор существующих программных средств и технологий для создания баз знаний онтологического типа;
- Задача 1-2.** анализ основных особенностей интеграции данных в онтологии и таксономия современных подходов к интеграции гетерогенных данных;
- Задача 1-3.** анализ возможностей и требований, предъявляемых к языкам описания отображения (mapping description languages), а также их сравнительный анализ;

Задача 1-4. обзор проблем, возникающих при интеграции данных на различных этапах жизненного цикла онтологии.

Раздел 2. Разработка методологии МЕТЕОР, включающей:

Задача 2-1. Выбор инструментов для работы с корпоративными онтологиями;

Задача 2-2. Анализ и сравнение основных методов формирования и наполнения онтологий;

Задача 2-3. Обоснование методологии формирования онтологий на основе интеграции данных из гетерогенных источников в онтологию с учетом когнитивных аспектов.

Все задачи первого этапа проекта МЕТЕОР выполнены.

Результаты работы по первому этапу проекта за 2017 год нашли отражение в 9 научных публикациях.

Заявленные цели на 2017 год

Применение корпоративных онтологий на предприятиях и в организациях создает потенциал для значительного повышения качества информационной поддержки и эффективности управления. Онтология (как термин Томаса Грубера из computer science, а не в философском смысле) представляет собой спецификацию концептуальной модели сложного объекта [Gruber, 1993]. Однако в настоящий момент слаба взаимосвязь между потребностями предприятий и организаций и существующими технологиями в области инженерии знаний и онтологического инжиниринга. Методы и инструменты работы с корпоративными знаниями пока недостаточно зрелы для решения практических задач управления знаниями и информационного менеджмента. В результате предприятия и организации применяют устаревшие и неэффективные технологии.

Проект МЕТЕОР ориентирован на решение актуальной проблемы разработки научных основ методологии и технологии формирования и наполнения онтологических баз знаний предметных областей на основе интеграции структурированных данных из гетерогенных источников. Интеграция данных является инструментом наполнения онтологии. Акцент в проекте ставится на учете не только синтаксических структур и грамматик интегрируемых данных, но и на рассмотрении семантических особенностей предметных областей. Разрабатываемые алгоритмы и модели предполагается полностью интегрировать с мировыми стандартами в области языков описания онтологий и открытых данных [<https://www.w3.org/standards/semanticweb/data>; <https://www.w3.org/OWL/>].

Цель проекта МЕТЕОР — разработка методологии и технологии, позволяющих создавать, наполнять и поддерживать базы знаний предприятий и организаций. Предполагается разработать технологию работы с гетерогенными, т.е. неоднородными по структуре и характеру, корпоративными данными и знаниями, а также архитектуру демонстрационного программно-методического прототипа решения МЕТЕОР-демо, интегрирующего отдельные потребности предприятий и организаций с научными разработками и технологиями в области онтологического инжиниринга и корпоративных онтологий. Реализация проекта позволит заполнить пробел между практическими потребностями предприятий и организаций и существующими технологиями и методами работы с онтологиями.

Целью первого этапа является анализ существующих подходов к формированию корпоративных онтологических баз знаний и к интеграции имеющихся данных в созданную онтологию, а также обзор, включающий:

- анализ основных особенностей интеграции данных в онтологии ;
- обзор существующих программных средств и технологий для создания баз знаний онтологического типа;
- таксономию современных подходов к интеграции гетерогенных данных;
- обзор проблем, возникающих при интеграции данных на различных этапах жизненного цикла онтологии (от проектирования до эксплуатации);
- возможности и требования, предъявляемые к языкам описания отображения (mapping description languages), а также их сравнительный анализ.

К целям первого этапа также относится разработка методологии формирования корпоративных баз знаний на основе онтологий, включающая:

- выбор инструментов для работы с корпоративными онтологиями;
- описание основных методов формирования и наполнения онтологий;
- обоснование метода интеграции данных в онтологию с учетом когнитивных аспектов.

Полученные за 2017 важнейшие научные результаты

Первый этап был посвящен анализу и изучению мирового опыта в области формирования онтологических баз знаний (БЗ) и интеграции данных из гетерогенных источников в БЗ предметных областей — корпоративных, научных, учебных и др. Под источниками данных понимаются реляционные БД, электронные таблицы и т.п.

Результаты по задачам раздела 1

Задача 1-1: обзор существующих программных средств и технологий для создания баз знаний онтологического типа

Активные публикации в области разработки и использования онтологий свидетельствуют о продолжающемся интересе и перспективности данной тематики [Buitelaar et al., 2006; Stucky, Luc, 2017; Mikhailov, Petrov, Lantow, 2016; Courtot, Malone, Mungall, 2016].

Онтологии позволяют интегрировать структурированную информацию, хранящуюся в корпоративных БД, с неструктурированной информацией по конкретной теме [Domingue, Fensel, Hendler, 2011]. Использование онтологий дает возможность значительно улучшить возможности управления знаниями в компаниях [Davies, 2013], что позволяет:

- приобретать знания (автоматическое извлечение информации);
- обеспечить их хранение в репозитории;
- обслуживать базу знаний (разработка, управление и поддержка онтологии);
- использовать знания (средства для поиска, обобщения, и визуализации знаний).

Разработка БЗ на онтологиях производится «сверху-вниз» (top-down), «снизу-вверх» (bottom-up) или «из центра в стороны» (middle-out) экземпляров [Pan et al, 2017]. Возможные варианты реализации указанных подходов:

1. В случае разработки «сверху-вниз» (от требований какие данные нужны):

- Формирования онтологии или сети онтологий
 - Наполнения онтологии или сети онтологий
2. В случае разработки «снизу-вверх» (от имеющихся данных):
 - Выявление онтологии/-й для источника/-ов данных
 - Разработка общей онтологии или создание сети онтологий
 - Установление соответствия между общей онтологией и онтологиями источников данных
 3. Комбинация первого и второго.

В современных системах используются два подхода в создании БЗ:

- начинают построение БЗ с общей онтологии (верхнего уровня) [Муромцев и др., 2011] и редактируют ее под требования конкретной БЗ
- при построении БЗ используется имеющиеся онтологии предметной области [Roldán-García, García-Nieto, Aldana-Montes, 2017; Viale et al., 2016].

Разработка больших систем (например, БЗ организации) связана с большой трудоемкостью создания единой онтологии [Тузовский, 2007], в связи с чем предлагаются две модели создания БЗ на основе онтологий:

- отказ от глубокой декомпозиции системы и включение в онтологию только наиболее значимых понятий из рассматриваемых предметных областей;
- выделение одного из направлений деятельности организации и создание для этого направления детальной, но узко специализированной онтологии.

При этом первый подход дает слишком грубую и обобщенную модель, а второй не позволяет использовать модель для взаимодействия между всеми подразделениями компании. Для решения этой проблемы предлагается создать иерархию областей знаний организации и дать возможность создавать отдельно онтологий разных подобластей, которые могут иметь разную детальность в зависимости от потребностей в их детальном моделировании.

Онтологические инструменты с актуальной поддержкой можно разделить на следующие функциональные группы, частично описанные в статье [Григорьев, Заблоцкий, Кудрявцев, 2012]:

1. Семантические Wiki: объединяют свойства вики-систем с семантическими утверждениями (Semantic MediaWiki);
2. Предметно-ориентированные инструменты (Enterprise Architect, ОПГ Мастер);
3. Онтологии для семантического поиска на сайте (Transinsight, Ubot);
4. Универсальные редакторы онтологий (Fluent Editor, Menthor Editor, owlged, Protégé, Semaphore Ontology Editor, Synaptica, Thesaurus Master, VocBench, Be Informed Suite);
5. Расширенные редакторы онтологий (Anzo, Information Workbench, Metaphacts, OntoStudio, ontotext, opensemanticframework, Poolparty, TopBraid).

Для реализации БЗ хорошо зарекомендовали себя как универсальные, так и расширенные редакторы онтологий, в то время как как остальные группы решают лишь специализированные задачи. Все перечисленные инструменты соответствуют стандартам W3C.

Обзор показал, что классическая палитра редакторов онтологий, широко описанная в литературе (PROTÉGÉ, ONTOedit и др.) сейчас активно пополняется новыми средствами (например, Fluent Editor), расширяющими и дополняющими существующие системы.

Основной результат обзора позволяет сделать вывод о том, что только инструменты двух групп (универсальные и расширенные редакторы онтологий) предназначены для работы с корпоративными онтологиями и для решения поставленных в проекте задач.

Задача 1-2: анализ основных особенностей интеграции данных в онтологии и таксономия современных подходов к интеграции гетерогенных данных

Классическая онтология ориентирована на моделирование предметных областей в понятиях классов и подклассов. По завершении моделирования классов, атрибутов, отношений и иерархии, необходимо наполнить онтологию экземплярами. Во многих крупных организациях данные или знания могут принимать различные форматы, такие как реляционные БД, веб-страницы, документация, и т.д. Чтобы сделать эту информацию доступной для общей онтологии, требуется преобразовать их из своего текущего представления в формат представления знаний.

Наполнение онтологий возможно в ручном и автоматизированном режимах. Варианты ручного наполнения БЗ рассмотрены в [Григорьев, Заблоцкий, Кудрявцев, 2012]. Для автоматизированного наполнения с точки зрения источников можно выделить:

- Использование структурированных данных (БД, таблицы, xml-файлы...) [Song, Zacharewicz, Chen, 2013];
- Использование неструктурированных данных (тексты). Применяется анализ текстов на естественном языке. Основные термины/подходы — семантическое аннотирование (semantic annotation), извлечение информации (information extraction) [Domingue, Fensel, Hendler, 2011].

Подходы к решению проблемы наполнения и обогащения онтологии зависят от уровня интеграции, свойств источников данных и требуемых способов и методов интеграции. Для структурированных данных эти подходы можно разделить на классы:

- по способу реализации интеграции:
 - отображение по требованию (mapping on-demand).
 - включение данных (консолидация) в базу знаний в качестве значений свойств объектов (data materialization).
- по использованию семантики при интеграции:
 - прямая интеграция (Direct Mapping) .
 - интеграция, учитывающая семантику предметной области (Domain Semantics-Driven Mappings) .

Таксономия современных подходов к интеграции гетерогенных структурированных данных представлена на рисунке 1.

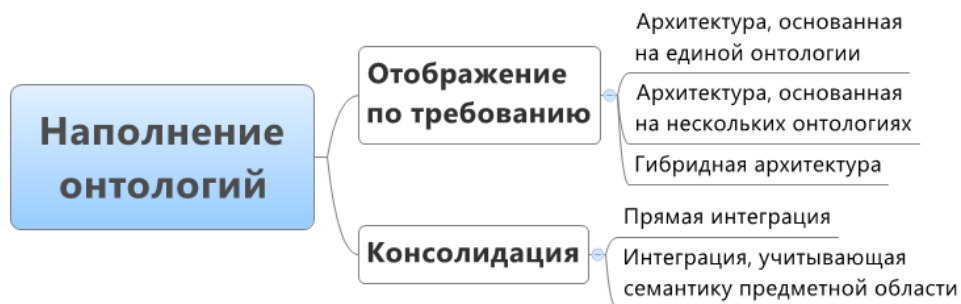


Рис. 1. Таксономия подходов к интеграции данных

Методы наполнения из структурированных источников данных варьируются в зависимости от типов данных и, как правило, основаны на их отображении (mapping) в онтологию. Способы извлечения информации для создания экземпляров [Pan et al., 2017]:

- Для CSV (значения, разделенные запятыми) и электронных таблиц есть расширение RDF для Google Refine, XLWrap, RDF123 и NOR₂O.
- Для XML: GRDDL через XSLT. TopBraid Composer и ReDeFer.
- Для других форматов any23 и Stats2RDF.

Большое количество информации содержится в реляционных БД. Для них создание и сопоставлений RDB2RDF вручную — трудоемкая задача, требующая значительных усилий и знаний [Pinkel et al., 2017]. В работе [Барахнин, Пастушков, 2015] описан подход, осуществляющий автоматический анализ определенных текстов для составления RDF. Однако можно воспользоваться готовыми решениями. Существует несколько инструментов, обеспечивающих технологическую поддержку этой задачи, например, D2R Server, ODEMapster, Triplify, Virtuoso RDF View и Ultrawrap.

Существующие специализированные программы, как например IncMap [Pinkel et al., 2017] могут обнаруживать и использовать семантические шаблоны в БД, а также использовать правила рассуждений в онтологии для преодоления структурных отличий.

Для более корректного процесса преобразования данных из реляционных систем в работе [Sitharamulu, Babu, 2016] описывается подход к преодолению недостатков для эффективного RDB сопоставления схем OWL / RDF. Предлагаемый подход известен как RDB-to-RDF Ingression. D2RI может эффективно сопоставлять данные двунаправленным образом от RDB до собственного хранилища RDF и наоборот.

Таким образом, рассмотрены основные особенности интеграции данных в онтологии и предложена новая классификация методов наполнения онтологий, позволяющая разработчикам осознанно выбирать и использовать наиболее подходящий способ в зависимости от стоящей перед ними задачи.

Задача 1-3: анализ возможностей и требований, предъявляемых к языкам описания отображения (mapping description languages), а также их сравнительный анализ

При преобразовании схемы БД данных в онтологию с учетом семантики предметной области правила отображения задаются явным образом с помощью языков описания отображения (mapping description languages). Эти языки можно разделить на две группы. Первая группа опирается на SQL запросы для описания отображения данных, что потенциально является недостатком, так как сложные случаи отображения не могут быть описаны с их помощью. С другой стороны, популярность SQL облегчает принятие таких подходов. Вторая группа использует специализированные языки описания отображения, что позволяет их создавать/расширять таким образом, чтобы удовлетворить любые специфические потребности. Но на практике выразительные возможности языков второй группы пока весьма ограничены [Michel, Montagnat, Faron-Zucker, 2014].

Были рассмотрены основные возможности, которые должны обеспечивать языки отображения и инструменты, их реализующие:

1. Генерация уникальных идентификаторов, определенных пользователем.
2. Наличие логических таблиц и выбор полей модели данных.
3. Переименование полей и возможности для условий выбора.
4. Использование существующей онтологии (возможность отображения реляционных объектов в экземпляры существующих онтологий)

5. Отображение одной таблицы в несколько классов
6. Преобразование отношения многие-ко-многим в простые «триады»
7. Создание анонимных узлов и трансляция типов данных
8. Возможность применения обработки данных.

Был проведен хронологический анализ истории языков описания отображений от языка DR2 MAP [Bizer, 2003], основанного на XML. С помощью языка DR2 MAP можно описать только простые правила отображения, поэтому потребовалось его развитие, результатом которого стал язык DR2Q [Cyganiak et al., 2012]. Далее язык D2R был трансформирован в R2O [Barrasa, Corcho, Gómez-Pérez, 2004]. Язык использовался в проектах ODEMapster [<http://neon-toolkit.org/wiki/ODEMapster>] и DB2OWL [Cullot, Ghawi, Yetongnon, 2007]. Также были рассмотрены проекты, использовавшие собственные языки отображения, например, METAmorphoses [Svihla, Jelinek, 2004], RDBToOnto [Cerbah, 2008], Relational.OWL [<http://sourceforge.net/projects/relational-owl/>; Laborda, Conrad, 2006], но их поддержка в настоящее время не осуществляется.

В 2012 году рабочая группа W3C RDB2RDF выпустила рекомендацию R2RML: RDB to RDF Mapping Language [<https://www.w3.org/TR/r2rml/>] (R2RML: Язык отображения RDB в RDF). Этот документ определяет язык описания отображения из реляционной базы данных в RDF, но не выдвигает никаких рекомендаций по реализации R2RML-процессора, поэтому существует ряд R2RML-совместимых инструментов с оригинальными подходами к реализации, например, Ultrawrap (<http://capsenta.com/ultrawrap>), DB2Triples (<https://github.com/antidot/db2triples>). R2RML вобрал в себя все лучшее от своих предшественников. Поэтому кажется, что использование R2RML неизбежно, когда речь идет о преобразовании реляционных данных в RDF.

Задача 1-4: обзор проблем, возникающих при интеграции данных на различных этапах жизненного цикла онтологии

Жизненный цикл онтологий может быть рассмотрен двояко. В широком смысле он охватывает этапы от появления идеи до реализации онтологии и её распространения [Farquhar et al., 1995; Fernandez, Gómez-Pérez, Juristo, 1997; Gandon, 2002; Li, Raskin, Ramani, 2007]. В более узком рассматривается как непосредственная разработка онтологии [Uschold et al., 1998; Noy, McGuinness, 2001; Nanda et al., 2006]. Попытка согласовать эти видения была сделана авторами методологии NeOn [Suárez-Figueroa et al., 2012], которые жизненный цикл понимают в широком смысле.

В модели жизненного цикла в широком смысле выделяются несколько общих стадий, на каждой из которых могут возникнуть проблемы наполнения онтологии экземплярами. Ряд таких проблем описан в [Petasis et al, 2011; Celjuska and Vargas-Vera, 2004]. В проекте были рассмотрены проблемы по этапам жизненного цикла онтологии:

1. Спецификация онтологии — определение цели построения и границ содержания. На этом же этапе должны быть решены вопросы включения в онтологию «чувствительных» (например, персональных) данных.
2. Разработка онтологии, в которую входят: сбор информации; концептуализация — формирование классов, свойств, их иерархии и отношений и др. Так как на этом этапе производится первичное наполнение онтологии данными, наибольшее количество проблем выявляется именно на нём. Среди них:

- a) распознавание объектов — определение частей данных, которые являются экземплярами;
 - b) соотнесение объектов с классами;
 - c) синонимия объектов — наличие разных значений, относящихся к одному объекту;
 - d) омонимия объектов — когда похожие объекты относятся к разным классам.
3. Оценка — проверка онтологии перед использованием. В случае автоматического или полуавтоматического наполнения должна быть проверена адекватность экземпляров категориям, их уникальность, согласованность, полнота и отсутствие избыточности.
 4. Поддержка в процессе использования. На этом же этапе может происходить эволюция — развитие онтологии, в том числе наполнение новыми терминами и обогащение. Развитие онтологии порождает проблемы обеспечения согласованности новых данных со старыми (от синтаксических, например, изменения формата ввода до семантических и структурных).
 5. Документирование, в том числе контроль изменений и взаимного соответствия версий. В процессе документирования должна описывается процедура наполнения онтологии и использованные приёмы.

Таким образом, рассмотрены основные проблемы интеграции данных на различных этапах жизненного цикла онтологии. Проект МЕТЕОР ориентирован на решении проблем 2 и 4 этапов, относящихся к наполнению БЗ экземплярами.

Результаты по задачам раздела 2

Задача 2-1: Выбор инструментов для работы с корпоративными онтологиями

Для подбора инструмента для работы с корпоративными онтологиями необходимо учитывать не только характеристики инструмента, но также особенности организации и класс задач, для решения которых планируется использовать онтологии.

- **Корпоративные задачи.** Онтологии могут использоваться для решения ряда задач, стоящих перед современной организацией: преодоления проблем коммуникации и совместного использования (sharing) знаний [Roche, 2000]; организации корпоративных веб-приложений [Obrst, Liu, Wray, 2003]; управления корпоративной памятью [Dieng et al., 1999]; организации бизнес-процессов [Hepp, Roman, 2007]; создания обучающих систем и т.д. Кроме этого, важно, какие методы работы с онтологиями будут использоваться в рамках этих задач, среди них: научение (ontology learning), оценка (ontology evaluation), объединение (ontology merging) и другие (см., например, [Corcho, Fernández-López, Gómez-Pérez, 2003]).
- **Характеристики компании.** При этом особенности использования онтологий в рамках одной задачи могут различаться в зависимости от характеристик компании, таких как: этап развития организации (уровень зрелости управления); масштаб (размер, степень глобализации); сектор (например, научно-исследовательский или производственный); производимый продукт; организационная структура (функциональная, проектная, командная и т.д.).

И корпоративные задачи, и характеристики компании влияют на выбор инструмента для работы с онтологиями. В отчете рассмотрены основные подходы к оценке инструментов и предложены рекомендации для подбора таких инструментов.

Существует несколько подходов к оценке инструментов по работе с онтологиями [Lambrix, Habbouche, Perez, 2003; Gómez-Pérez, 2004; Murshed, Singh, 2005; Duineveld et al., 2000; Corcho, Fernández-López, Gómez-Pérez, 2003].

В проекте предлагаемые подходы расширены и адаптированы к специфике работы с корпоративными онтологиями. Предлагаемая методика включает два набора параметров: общие, относящиеся к «оболочке» инструмента и функциональные, описывающие непосредственные возможности для построения онтологий. Часть параметров носит «абсолютный характер» - например, если инструмент не соответствует стандартам Консорциума всемирной паутины, его использование не рекомендуется. Такие параметры помечены астериском (*). Большинство параметров носят «относительный» характер и должны быть учтены в соответствии с конкретными целями и задачами компании. Для таких параметров впервые предложены рекомендации по выбору.

Общие параметры:

1. *Соответствие стандартам World Wide Web Consortium. Более того, для отдельных областей могут существовать собственные стандарты, например, BioOntology Consortium рекомендует DAML+OIL для биологических онтологий.
2. Планируемая методология разработки. Описано множество подходов к построению онтологий [Fernández-López, 1999], выбирать инструмент следует исходя из того, какой подход планирует использовать организация.

Функциональные параметры:

1. Технические:
 - 1.1. *Совместимость, в частности, возможность экспорта/импорта онтологии в общепринятом формате (OWL, RDF).
 - 1.2. Реализация и скорость навигации по онтологии, в том числе наличие семантического поиска.
 - 1.3. Языки описания онтологий и представления аксиом должны быть учтены в случае, если планируется работа с исходными форматами данных.
 - 1.4. Поддержка визуализации.
 - 1.5. Встроенный аналитический модуль.
2. Возможности моделирования:
 - 2.1. *Наличие модуля валидации и верификации для проверки согласованности онтологии и валидности языка описания, как в процессе разработки, так и в процессе обогащения (при добавлении новых классов и свойств).
 - 2.2. Реализация логического вывода. Может быть как частью инструмента, так и отдельным подключаемым модулем.
 - 2.3. Уровень реализации автоматического извлечения информации. Важен при потребности реализации scratch-методов, таких как машинное обучение, обработка естественного языка.

- 2.4. Поддерживаемый уровень экспрессивности должен подбираться под конкретные задачи компании, например, множественное наследование может как быть, так и не быть критически важным параметром.
- 2.5. Примеры реализации распространённых шаблонов проектирования онтологий (Ontology Design Patterns).
- 2.6. Формирование схемы на основе структуры источников данных.
- 2.7. Возможности работы с несколькими онтологиями (например, слияния).
- 2.8. Возможность работы с общедоступными онтологиями и словарями (например, DBpedia)
3. Возможности наполнения экземплярами
 - 3.1. Наполнение онтологии экземплярами из разных источников (баз данных, таблиц, полу- и неструктурированных файлов), в том числе в автоматическом или полуавтоматическом режиме.
 - 3.2. Реализация модификации и удаления экземпляров.

Данные параметры являются критически важными, в случае, если организация планирует наполнять и обогащать онтологию (например, при создании онтологической базы знаний) и не важны, в случае, если онтология используется только для моделирования.

4. Принадлежность к определённой функциональной группе (не является критически важным параметром, но может помочь при выборе инструмента под определённые задачи).
 - 4.1. Семантические Wiki подойдут для расширения функционала базы знаний, реализованной на классическом Wiki-движке.
 - 4.2. Предметно-ориентированные инструменты.
 - 4.3. Онтологии для семантического поиска на сайте для веб-приложений.
 - 4.4. Редакторы онтологий (универсальные и расширенные) для разработки онтологии, их редактирования, анализа и наполнения экземплярами.

Рассмотренные характеристики могут быть использованы при выборе инструмента для работы с корпоративными базами знаний онтологического типа, при этом в зависимости от конкретных задач и характеристик организации разным параметрам должен быть присвоен разный вес.

Таким образом, предложена новая методика выбора инструментов для работы с корпоративными онтологиями, учитывающая не только характеристики инструмента, но также особенности организации, в которой инструмент будет внедряться, и класс задач, для решения которых планируется использовать онтологии.

Задача 2-2: Анализ и сравнение основных методов формирования и наполнения онтологий

Разработка базы знаний онтологического типа включает [Pan et al, 2017]:

- Выявление требований к создаваемой системе и её базе знаний.
- Формирования онтологии или сети онтологий.
- Наполнения онтологии или сети онтологий.
- Публикация полученной базы для дальнейшего использования.

Основные последние методологии разработки онтологий рассмотрены в работе [Гаврилова и др., 2016; Simperl, Luczak-Rösch, 2014]: DOGMA-MESS [De Moor, De Leenheer, Meersman, 2006; De Leenheer, Debruyne, 2008], DILIGENT [Tempich et al, 2005,

2007], NeOn (www.neon-project.org) [Suárez-Figueroa, Gómez-Pérez, Fernández-López, 2012] и др. Важнейшая современная рекомендация по разработке онтологий - использовать готовыми доступные онтологии, описывающими предметную область, которые можно найти в Swoogle (<http://swoogle.umbc.edu/>) или в LOV (Linked Open Vocabularies <http://lov.okfn.org/>), либо максимально включить их в процесс создания онтологии [Suárez-Figueroa, Gómez-Pérez, Fernández-López, 2012].

Также возможна интеграция нескольких онтологий для получения единого информационного пространства [Бова, 2015], например, на основе сравнения онтологических схем и интеграции концептов, с последующим формированием шаблона запросов для взаимного отображения онтологий друг в друга.

Проведенный обзор и пилотные исследования показали, что существующие методы наполнения и обогащения баз знаний онтологического типа обладают рядом недостатков:

- Необходимость использования и, соответственно, изучения синтаксиса нескольких специфических языков для описания отображений.
- Использование различных методов для каждого типа источника данных.
- Отсутствие готовых решений, обеспечивающих одновременную интеграцию данные из нескольких источников.
- Отсутствие рекомендаций по устранению противоречий в данных из различных источников.
- Компании могут использовать проприетарные форматы хранения данных и в будущем могут появиться новые. Рекомендации по интеграции данных в подобном случае отсутствуют.

Задача 2-3: Обоснование метода МЕТЕОР (интеграции данных) в онтологию с учетом когнитивных аспектов

В проекте предлагается метод автоматизированного наполнения и обогащения онтологий или БЗ онтологического типа путем консолидации на основе структурированных данных, хранящихся в различных гетерогенных источниках. Основными типами источников структурированных данных являются реляционные базы данных, XML-документы, электронные таблицы, а также текстовые файлы с проприетарной структурой. Алгоритм процесса консолидации данных представлен на рисунке 2.

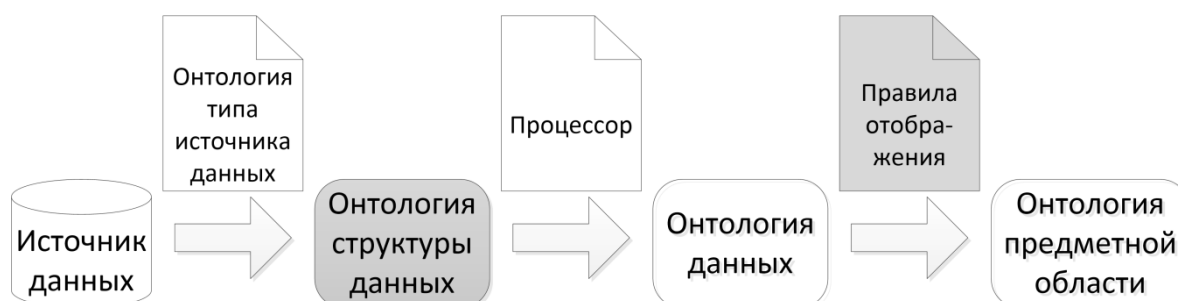


Рис. 2. Алгоритм процесса консолидации данных

Для возможности обработки данных из источников различных типов необходимо создать онтологии, описывающие структуру этих источников в общем виде. Например, онтология реляционной базы данных в упрощенном виде представлена на рисунке 3.

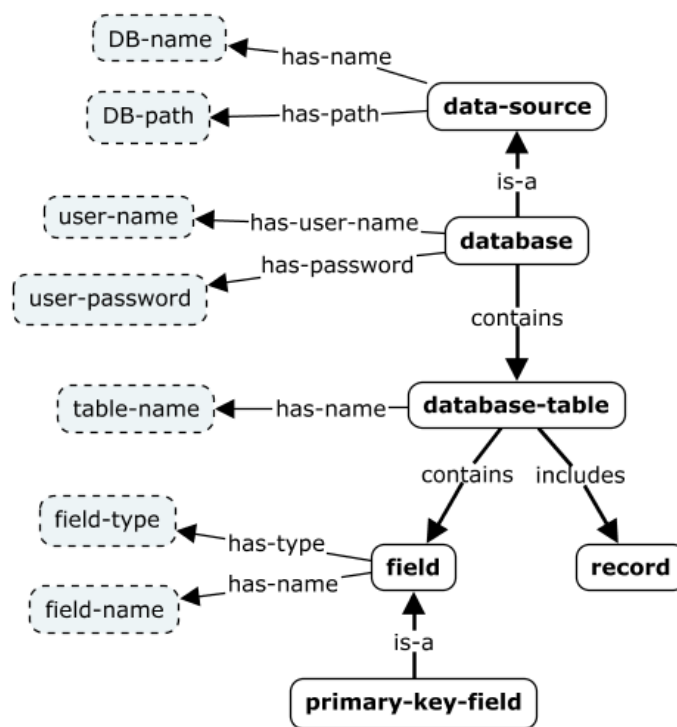


Рис. 3. Онтология реляционной базы данных

Используя онтологию типа источника данных, разработчик описывает структуру источника, данные из которого необходимо поместить в БЗ. Специальная программа-процессор, используя информацию из онтологии структуры данных, формирует промежуточную онтологию, содержащую данные из описанного источника.

С помощью правил отображения разработчик описывает, как именно данные будут интегрированы в онтологию. Правила также могут регламентировать поведение процесса интеграции в случае возникновения противоречий в данных из различных источников. В результате работы правил, онтология предметной области обогащается с использованием данных из промежуточных онтологий.

Предлагаемый метод имеет модульную архитектуру и потенциально может использоваться для консолидации структурированных данных из источников любых типов. Он позволяет интегрировать данные из источников различного типа, учитывает их распределенную природу, необходимость аутентификации при доступе к сетевым ресурсам, а также предлагает решение проблемы разрешения противоречий в интегрируемых данных, что позволит снизить трудоемкость наполнения и обогащения онтологий, используя накопленные массивы информации независимо от формы их хранения и представления.

Таким образом, предложен новый метод наполнения онтологий структурированными данными, хранящихся в различных гетерогенных источниках. Метод лишен большинства указанных выше недостатков. Предлагаемое решение является расширяемым для обеспечения потенциальной возможности работы с новыми форматами данных.

Степень новизны полученных результатов

В России активно ведутся работы в области интеграции данных с использованием онтологий (коллективы под руководством В.Ф. Тузовского, Л.А. Калиниченко,

М.Р. Когаловского и др.). Однако в этих исследованиях ставится иная задача: необходимо интегрировать данные из разнородных источников в информационную систему, предоставляющую пользователю единый интерфейс доступа к этим данным. Онтология в этом контексте выступает в качестве инструмента, а не результата. Задача интеграции данных в онтологию предметной области с учетом когнитивных аспектов ставится в России впервые.

Предлагаемый метод МЕТЕОР является оригинальным новым методом, позволяющим гибко настраивать интеграцию данных из источников различного типа в ранее созданную онтологию; поддерживающим модульность и учитывающим проблему устаревания данных.

Идея сопряжения технических и функциональных характеристик с задачи и особенностями организации при выборе инструментов для работы с корпоративными онтологиями также является новой.

Поэтому результаты первого этапа проекта МЕТЕОР отличаются оригинальностью и новизной и предлагают идеи и методы, востребованные во многих компаниях, доросших до важности проблем УЗ и готовых к внедрению корпоративных баз знаний (РОСНАНО, РосАТОМ, ряд компаний сектора ИКТ).

Сопоставление полученных результатов с мировым уровнем

Онтологии являются одним из наиболее перспективных средств представления знаний. Применение корпоративных онтологий на предприятиях и в организациях создает потенциал для значительного повышения качества информационной поддержки и эффективности управления. Однако в настоящий момент узким местом является автоматизированное расширение (наполнение и обогащение) существующей онтологии, как правило, на основе информации, извлеченной из предметно-ориентированного контента. Уже 15 лет в мире не ослабевает интерес к работам, посвященным наполнению и обогащению онтологий [Valarakos et al., 2004; Petasis et al., 2011; Nederstigt et al., 2014]. К недостаткам существующих подходов можно отнести их узкую направленность на решение задач конкретного проекта, обеспечение возможности интеграции данных из источников только одного типа, отсутствие модульности и расширяемости, а также удобного инструментария. Разрабатываемая методика наполнения и обогащения онтологий направлена на устранение вышеуказанных сдерживающих факторов и предназначена для использования в корпоративных онтологиях.

Разработка метода формирования онтологий и интеграции структурированных данных из гетерогенных источников МЕТЕОР основывается на существующих технологиях интеграции, учитывающих семантику предметной области (Domain Semantics-Driven Mappings), использующих языки описания отображений (mapping description languages [Cyganiak et al., 2012; <https://www.w3.org/TR/rdb-direct-mapping/>]). Предлагаемый метод основывается на обобщении экспертных знаний и мирового опыта, например, [Song, Zacharewicz, Chen, 2013; Pinkel et al, 2017; Pan et al, 2017; Sitharamulu, Babu, 2016; Michel, Montagnat, Faron-Zucker, 2014; Suárez-Figueroa, Gómez-Pérez, Fernández-López, 2012; Paulheim, 2017; Бова, 2015], и принимает во внимание преимущества и недостатки предыдущих технологий интеграции.

В целом, разрабатываемый метод МЕТЕОР, с одной стороны, находится в русле мировых трендов, с другой стороны, является оригинальным новым методом,

позволяющим гибко настраивать интеграцию данных из источников различного типа в ранее созданную онтологию; поддерживающим модульность и учитывающим проблему устаревания данных. Эта методология позволит проектировать и поддерживать корпоративные базы знаний для эффективного информационного менеджмента.

Также одним из результатов первого этапа проекта является создание методики выбора инструментов для работы с корпоративными онтологиями. В качестве критериев выбора инструмента в предлагаемой методике используются не только характеристики инструмента, но и особенности организации, в которой инструмент будет внедряться, а также классы задач, для решения которых планируется использовать онтологии. Предлагаемая методика обобщает и расширяет результаты, полученные в [Lambrix, Habbouche, Perez, 2003; Gómez-Pérez, 2004; Murshed, Singh, 2005; Duineveld et al., 2000; Corcho, Fernández-López, Gómez-Pérez, 2003]. Таким образом, результаты по задаче 2-1 продолжают мировой тренд исследования применимости современных инструментов для создания и поддержки корпоративных онтологий с учетом их специфики.

Успешные выступления авторов на российских и европейских конференциях 2017 года в области управления знаниями и интеллектуальных систем (например, пленарный доклад «Новые методологии формирования онтологий» на VII-й Всероссийской научно-практической конференции «Нечеткие системы, мягкие вычисления и интеллектуальные технологии» и доклад на VII-й международной конференции «Системный анализ и информационные технологии») также подтверждают международный уровень представленных в работе результатов первого этапа.

Методы и подходы, использованные в ходе выполнения проекта

Первый этап проекта МЕТЕОР был посвящен анализу проблематики и состояния исследований в области формирования корпоративных онтологических баз знаний и интеграции в них данных из гетерогенных источников.

Задачи первого этапа направлены на широкое аналитическое исследование state-of-the-art в области онтологического инжиниринга в целом и создания/наполнения онтологических баз знаний в частности. Для проведения этого исследования использовались следующие методы:

1. системный анализ;
2. критический обзор литературы, технологий, платформ и сервисов (более 50 источников);
3. структурирование информации через формирование интеллект-карт;
4. методы онтологического инжиниринга;
5. контекстный анализ;
6. инструменты Google Analytics;
7. обобщение, категоризация и синтез понятий для построения таксономий.

В таком сочетании эти методы использованы впервые, а их синергетический эффект позволил получить новые оригинальные результаты, опубликованные в 9 научных работах.

Разработка нового метода формирования онтологий и интеграции структурированных данных из гетерогенных источников с учетом когнитивных аспектов была основана на системном подходе с использованием принципов онтологического инжиниринга и на существующих технологиях интеграции, учитывающих семантику

предметной области (Domain Semantics-Driven Mappings), использующих языки описания отображений (mapping description languages). Предлагаемые метод принимает во внимание преимущества и недостатки предыдущих технологий интеграции, учтенные, в частности, в рекомендациях рабочей группы W3C-консорциума.

План работ на 2018 год

Второй этап по созданию новой технологии формирования баз знаний на основе онтологий будет основываться на современных подходах к интеграции, учитывающих семантику предметной области (Domain Semantics-Driven Mappings). Он будет посвящен разработке новой технологии, позволяющей компаниям эффективно создавать базы знаний и затем интегрировать в них ранее накопленную информацию и данные. Подход будет учитывать особенности жизненного цикла онтологий, в частности, проблему устаревания данных. Также будет предложена типовая архитектура системы.

Задачи 2-ого этапа:

1. Создание технологии МЕТЕОР формирования онтологий и интеграции структурированных данных из гетерогенных источников с учетом когнитивных аспектов:
 - формальное описание технологии формирования и наполнения онтологий (Гаврилова Т.А., Лещева И.А., Кудрявцев Д.В.)
 - создание онтологий источников данных различных типов (реляционная база данных, XML-документ, электронная таблица MS Excel); (Гаврилова Т.А., Лещева И.А., Беглер А.М., Кубельский М.В.)
 - разработка рекомендаций по описанию источников данных с использованием созданных онтологий; (Плешкова А.Ю., Благов Е.Ю., Страхович Э.В.)
2. Разработка архитектура системы, поддерживающей/реализующей технологию МЕТЕОР; (Лещева И.А., Кудрявцев Д.В., Кубельский М.В.)
3. Разработка пилотного прототипа программного комплекса, реализующего технологию МЕТЕОР, включающая:
 - разработку препроцессора, формирующего промежуточную онтологию, содержащую данные из гетерогенных источников; (Лещева И.А.)
 - формирование рекомендаций по использованию созданного прототипа (Беглер А.М.).

Библиографический список всех публикаций по проекту за 2017 год

1. Гаврилова Т.А., Онуфриев В. А. Анализ ошибок студентов при визуальном структурировании знаний // Компьютерные инструменты в образовании, № 6, 2016. – с. 41-53.
2. Гаврилова Т.А., Страхович Э.В. Системный подход к разработке интеллект-карт // Сборник научных трудов XXI международной научно-практической конференции "Системный анализ в проектировании и управлении (SAEC-2017)", Санкт-Петербург, 29-30 июня 2017. – с. 226-234.
3. Лещева И.А. Формирование баз знаний через наполнение и обогащение онтологий // Сборник научных трудов 20-ой Российской научно-практической конференции

- "Инжиниринг предприятий и управление знаниями (ИП&УЗ-2017)". Моск. госуд. ун-т экономики, статистики и информатики. – М., 2017. – с. 68-73.
4. Лещева И.А. Проблемы интеграции данных на различных этапах жизненного цикла онтологии // Труды 7-ой международной конференции «Системный анализ и информационные технологии (САИТ-2017)», Балтийский федеральный университет им. И. Канта, г. Светлогорск, Калининградская область, 13-18 июня 2017. – с. 279-284.
 5. Gavrilova T., Alsufyev A., Grinberg E., Mailov E. Three Historical Phases in Research on Visual Knowledge Models // Abstract Book of 11th International Forum on Knowledge Assets Dynamics (IFKAD-2017), St. Petersburg, Russia, June 7-9, 2017. – p. 102. Proceeding online: pp. 840-851.
 6. Gavrilova T., Alsufyev A., Kokoullina L. Knowledge Management in Russian Companies: Overall Score // Proc. of the International Conference «Theory and Applications in the Knowledge Economy (TAKE 2017)», Zagreb, Croatia, July 12-14, 2017. – pp. 637-645.
 7. Gavrilova T., Alsufyev A., Grinberg E., Mailov E. Knowledge Visualization as a new Discipline: Swiss Bias // Conference Proceedings of 18th European Conference on Knowledge Management (ECKM-2017), Barcelona, September 6-7, 2017. – pp. 345-352.
 8. Leshcheva I., Blagov E., Pleshkova A. Towards a Method of Ontology Population from Heterogeneous Sources of Structured Data // Proc. of 30th Jubilee Neumann Colloquium, Budapest, Hungary, November 24-25, 2017. – pp. 29-34.
 9. Onufriev V., Gavrilova T. Conceptual Modelling: Common Students' Mistakes In Visual Representation // Proceeding of International Conference on Interactive Collaborative Learning (ICL2017) – Teaching and Learning in a Digital World, Budapest, Hungary, September 27-29, 2017. – pp. 62-71.

Список использованной литературы

1. Барахнин В. Б., Пастушков И. С. Технология автоматизированного наполнения онтологии фактографической поисковой системы //Вестник Новосибирского государственного университета. Серия: Информационные технологии. – 2015. – Т. 13. – №. 4. – с. 5-12.
2. Бова В. В. Онтологическая модель интеграции данных и знаний в интеллектуальных информационных системах //Известия Южного федерального университета. Технические науки. – 2015. – №. 4 (165). – с. 225-237.
3. Гаврилова Т. А., Кудрявцев Д. В., Муромцев Д. И. (2016). Инженерия знаний. Модели и методы: Учебник. //СПб.: Издательство Лань. – 2016. – 348 с.
4. Григорьев Л. Ю., Заблоцкий А. А., Кудрявцев Д. В. Технология наполнения баз знаний онтологического типа //Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Информатика. Телекоммуникации. Управление. – 2012. – №. 3 (150). – с. 27-36.
5. Муромцев Д. и др. Разработка базы знаний по оптике для образовательных веб-приложений //Искусственный интеллект и принятие решений. – 2011. – №. 3. – с. 3-10.
6. Тузовский А. Ф. Разработка систем управления знаниями на основе единой онтологической базы знаний //Известия Томского политехнического университета. – 2007. – Т. 310. – №. 2. – с. 182-185.

7. Barrasa J., Corcho Ó., Gómez-Pérez A. R2O, an Extensible and Semantically based Database-to-Ontology Mapping Language //Proceedings of the 2nd Workshop on Semantic Web and Databases (SWDB2004), Toronto, Canada, 2004. – pp. 1069-1070.
8. Bizer C. D2R MAP - a database to RDF mapping language [Electronic resource] //Proceedings of the 12th International World Wide Web Conference, Budapest, Hungary, 2003. URL: <http://wwwconference.org/www2003/cdrom/papers/poster/p004/p4-bizer.html>
9. Buitelaar P. et al. Ontology-based Information Extraction with SOBA [Electronic resource] //Proceedings of the International Conference on Language Resources and Evaluation (LREC-2006), Genoa, Italy, 2006. – pp. 2321–2324. URL: http://lrec-conf.org/proceedings/lrec2006/pdf/93_pdf.pdf
10. Celjaska D., Vargas-Vera M. Ontosophie: A semi-automatic system for ontology population from text // International Conference on Natural Language Processing (ICON). – 2004. – pp. 60-73.
11. Cerbah F. Learning highly structured semantic repositories from relational databases //The semantic web: Research and applications. ESWC 2008. Lecture Notes in Computer Science. Vol. 5021. Springer Berlin Heidelberg – 2008. – C. 777-781. doi: 10.1007/978-3-540-68234-9_57
12. Corcho O., Fernández-López M., Gómez-Pérez A. Methodologies, tools and languages for building ontologies. Where is their meeting point? // Data & knowledge engineering. – 2003. – Vol. 46. – №. 1. – pp. 41-64.
13. Courtot M., Malone J., Mungall C. Ten Simple Rules for Biomedical Ontology Development [Electronic resource] //Proceedings of the Joint International Conference on Biological Ontology and BioCreative, Corvallis, Oregon, United States, 2016. URL: http://ceur-ws.org/Vol-1747/IT404_ICBO2016.pdf
14. Cullot N., Ghawi R., Yetongnon K. DB2OWL : A tool for automatic database-to-ontology mapping //Proceedings of the 15th Italian Symposium on Advanced Database Systems, Torre Canne - Fasano, Brindisi, Italy, 2007. – pp. 491–494.
15. Cyganiak R. et al. The D2RQ mapping language v0.8, 2012. URL: <http://d2rq.org/d2rq-language>
16. Davies J., Fensel D., Van Harmelen F. (ed.). Towards the semantic web: ontology-driven knowledge management. – Chichester, England : John Wiley & Sons, 2003. 288 p.
17. De Laborda C. P., Conrad S. Database to semantic web mapping using RDF query languages //Proceedings of the 25th International Conference on Conceptual Modeling (ER 2006), Tucson, AZ, USA, 2006. Lecture Notes in Computer Science. Vol. 4215.– pp. 241-254.
18. De Leenheer P., Debruyne C. DOGMA-MESS: A tool for fact-oriented collaborative ontology evolution //OTM Confederated International Conferences "On the Move to Meaningful Internet Systems". – Springer, Berlin, Heidelberg, 2008. – pp. 797-806.
19. De Moor A., De Leenheer P., Meersman R. DOGMA-MESS: A meaning evolution support system for interorganizational ontology engineering //International Conference on Conceptual Structures. – Springer, Berlin, Heidelberg, 2006. – pp. 189-202.
20. Dieng R. et al. Methods and tools for corporate knowledge management // International journal of human-computer studies. – 1999. – Vol. 51. – №. 3. – pp. 567-598.
21. Domingue J., Fensel D., Hendler J. A. (ed.). Handbook of semantic web technologies. – Berlin, Heidelberg : Springer-Verlag Berlin Heidelberg, 2011. 1035 p. doi: 10.1007/978-3-540-92913-0

22. Duineveld A. J. et al. WonderTools? A comparative study of ontological engineering tools // *International Journal of Human-Computer Studies*. – 2000. – Vol. 52. – №. 6. – pp. 1111-1133.
23. Farquhar A. et al. Collaborative ontology construction for information integration. – Technical Report KSL-95-63, Stanford University Knowledge Systems Laboratory, 1995.
24. Fernández-López M. Overview of methodologies for building ontologies // *Proceedings of the IJCAI-99 workshop on Ontologies and Problem-Solving Methods (KRR5)*. – 1999.
25. Fernandez M., Gómez-Pérez A., Juristo N. Methontology: from ontological art towards ontological engineering // *Proceedings of the AAAI97 spring symposium series on ontological engineering*. – 1997. – pp. 33-40.
26. Gandon F. *Ontology Engineering: a survey and a return on experience* [Electronic resource] // INRIA, 2002. URL: <https://hal.inria.fr/inria-00072192/document>
27. Gómez-Pérez A. *Ontology evaluation* // *Handbook on ontologies*. – Springer Berlin Heidelberg, 2004. – pp. 251-273.
28. Hepp M., Roman D. An ontology framework for semantic business process management [Electronic resource] // *Wirtschaftsinformatik Proceedings 2007*. URL: <https://aisel.aisnet.org/wi2007/27>
29. Lambrix P., Habbouche M., Perez M. Evaluation of ontology development tools for bioinformatics // *Bioinformatics*. – 2003. – Vol. 19. – №. 12. – pp. 1564-1571.
30. Li Z., Raskin V., Ramani K. A methodology of engineering ontology development for information retrieval // *Proceedings of the 16th International Conference on Engineering Design (ICED'07)*. – 2007.
31. Michel F., Montagnat J., Faron-Zucker C. A survey of RDB to RDF translation approaches and tools: Research Report. – 2014. URL: <http://hal.archives-ouvertes.fr/hal-00903568>
32. Mikhailov S., Petrov M., Lantow B. *Ontology Visualization: A Systematic Literature Analysis* [Electronic resource] // *Joint Proceedings of the BIR 2016 Workshops and Doctoral Consortium co-located with 15th International Conference on Perspectives in Business Informatics Research (BIR 2016)*, Prague, Czech Republic, 2016. URL: <http://ceur-ws.org/Vol-1684/paper24.pdf>
33. Murshed M., Singh R. Evaluation and ranking of ontology construction tools [Electronic resource] // *Technical Report DIT-05-013*, University of Trento, 2005. URL: <http://eprints.biblio.unitn.it/747/1/013.pdf>
34. Nanda J. et al. A methodology for product family ontology development using formal concept analysis and web ontology language // *Journal of computing and information science in engineering*. – 2006. – Vol. 6. – №. 2. – pp. 103-113.
35. Niderstigt L. J. et al. Floppies: A framework for large-scale ontology population of product information from tabular data in e-commerce stores // *Decision Support Systems*. – 2014. – Vol. 59. – pp. 296-311.
36. Noy N. F., McGuinness D. L. *Ontology development 101: A guide to creating your first ontology* [Electronic resource]. – 2001. URL: ftp://ksl.stanford.edu/pub/KSL_Reports/KSL-01-05.pdf.gz
37. Obrst L., Liu H., Wray R. *Ontologies for corporate web applications* // *AI Magazine*. – 2003. – Vol. 24. – №. 3. – pp. 49-62.
38. Pan J. Z. et al. (ed.). *Exploiting Linked Data and Knowledge Graphs in Large Organisations*. – Cham : Springer, 2017. 266 p. doi: 10.1007/978-3-319-45654-6

39. Paulheim H. Knowledge graph refinement: A survey of approaches and evaluation methods //Semantic web. – 2017. – Vol. 8. – №. 3. – pp. 489-508.
40. Petasis G. et al. Ontology population and enrichment: State of the art // Knowledge-driven multimedia information extraction and ontology evolution. – Springer-Verlag, 2011. – pp. 134-166.
41. Pinkel C. et al. IncMap: A Journey towards Ontology-based Data Integration [Electronic resource] //Proceedings of the 17th Conference on Database Systems for Business, Technology, and Web (BTW-2017), Stuttgart, Germany, 2017. – pp. 145-164. URL: http://btw2017.informatik.uni-stuttgart.de/slidesandpapers/F3-10-15/paper_web.pdf
42. Roche C. Corporate ontologies and concurrent engineering // Journal of Materials Processing Technology. – 2000. – Vol. 107. – №. 1. – pp. 187-193.
43. Roldán-García M. D. M., García-Nieto J., Aldana-Montes J. F. Enhancing semantic consistency in anti-fraud rule-based expert systems //Expert Systems with Applications. – 2017. – Vol. 90. – pp. 332-343.
44. Simperl E., Luczak-Rösch M. Collaborative ontology engineering: a survey //The Knowledge Engineering Review. – 2014. – Vol. 29. – №. 1. – pp. 101-131.
45. Sitharamulu V., Babu B. R. A Novel Proposal for Bridging Gap between RDB-RDF Semantic Web using Bidirectional Approach //International Journal of Applied Engineering Research. – 2016. – Vol. 11. – №. 6. – pp. 4456-4460.
46. Song F., Zacharewicz G., Chen D. An ontology-driven framework towards building enterprise semantic information layer //Advanced Engineering Informatics. – 2013. – Vol. 27. – №. 1. – pp. 38-50.
47. Stucky B., Luc A. OntoPilot: New Software to Simplify and Accelerate Ontology Development and Deployment //Biodiversity Information Science and Standards. – 2017. – Vol. 1. – e20192. doi: 10.3897/tdwgproceedings.1.20192
48. Suárez-Figueroa M. C., Gómez-Pérez A., Fernández-López M. The NeOn methodology for ontology engineering //Ontology engineering in a networked world. – Springer Berlin Heidelberg, 2012. – pp. 9-34. doi: 10.1007/978-3-642-24794-1_2
49. Svihla M., Jelinek I. Two layer mapping from database to RDF //Proceedings of the 7th Electronic Computers and Informatics (ECI 2004), Košice-Herľany, Slovakia, 2004.
50. Tempich C. et al. An argumentation ontology for distributed, loosely-controlled and evolving engineering processes of ontologies (DILIGENT) //European Semantic Web Conference. – Springer, Berlin, Heidelberg, 2005. – pp. 241-256.
51. Tempich C. et al. Argumentation-based ontology engineering //IEEE Intelligent Systems. – 2007. – Vol. 22. – №. 6. – pp. 52–59.
52. Uschold M. et al. The enterprise ontology //The knowledge engineering review. – 1998. – Vol. 13. – №. 1. – pp. 31-89.
53. Valarakos A. G. et al. Enhancing ontological knowledge through ontology population and enrichment //International Conference on Knowledge Engineering and Knowledge Management. – Springer, Berlin, Heidelberg, 2004. – pp. 144-156.
54. Viale P. et al. Human endocrine system modeling based on ontologies //Knowledge-Based Systems. – 2016. – Vol. 111. – pp. 113-132.

Участие в научных мероприятиях по тематике проекта за 2017 год

1. VII-я Всероссийская научно-практическая конференция «Нечеткие системы, мягкие вычисления и интеллектуальные технологии (НСМВИТ-2017)», пленарный доклад, Гаврилова Т.А., Лещева И.А. «Методологии формирования онтологий: краткий обзор»
2. XXI Международная научно-практическая конференция «Системный анализ в проектировании и управлении (SAEC-2017)», доклад на секции, Гаврилова Т.А., Страхович Э.В. «Системный подход к разработке интеллект-карт»
3. 20-я Российская научно-практическая конференция «Инжиниринг предприятий и управление знаниями (ИП&УЗ-2017)», доклад на секции, Лещева И.А. «Формирование баз знаний через наполнение и обогащение онтологий»
4. 7-я Международная конференция «Системный анализ и информационные технологии (САИТ-2017)», доклад, Лещева И.А. «Проблемы интеграции данных на различных этапах жизненного цикла онтологии»
5. 11th International Forum on Knowledge Assets Dynamics (IFKAD-2017), доклад на секции, Gavrilova T., Alsufyev A., Grinberg E., Mailov E. «Three Historical Phases in Research on Visual Knowledge Models»
6. The International Conference «Theory and Applications in the Knowledge Economy (TAKE 2017)», доклад на секции, Gavrilova T., Alsufyev A., Kokoullina L. «Knowledge Management in Russian Companies: Overall Score»
7. 18th European Conference on Knowledge Management (ECKM-2017), доклад на секции, Gavrilova T., Alsufyev A., Grinberg E., Mailov E. «Knowledge Visualization as a new Discipline: Swiss Bias»
8. 30th Jubilee Neumann Colloquium, доклад, Leshcheva I., Blagov E., Pleshkova A. «Towards a Method of Ontology Population from Heterogeneous Sources of Structured Data»
9. Interntional Conference on Interactive Collaborative Learning (ICL2017), доклад на секции, Onufriev V., Gavrilova T. «Conceptual Modelling: Common Students' Mistakes In Visual Representation»