

Итоговый Отчет за 2019 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

Проект МЕТЕОР был 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 состоял из трех этапов:

1. На первом этапе (2017) был проведен анализ состояния современных исследований в области формирования и поддержки баз знаний онтологического типа и разработаны основы методологии интеграции структурированных данных из гетерогенных источников в базы знаний различных предметных областей – корпоративных, научных, учебных и др.
2. На втором этапе (2018) была разработана МЕТЕОР-технологии формирования онтологий и интеграции структурированных данных из гетерогенных источников с учетом человеческого фактора, в частности, когнитивных аспектов дизайна и генерации онтологий.
3. Третий заключительный этап (2019) был посвящен разработке архитектуры прототипной системы МЕТЕОР-демо, поддерживающей/реализующей эту технологию и апробация разработанных методологии и технологии на примере создания базы знаний учебного процесса ВУЗа.

На третьем этапе были решены следующие задачи:

- Анализ административных задач, возникающих в рамках управления и обеспечения учебных процессов. Выбор задачи для кастомизации методологии создания базы знаний и ее архитектуры.
- Апробация разработанного подхода, которая включила в себя:
 - формирование пилотного прототипа базы знаний для решения выбранной задачи;
 - выбор источников данных для заполнения базы знаний;
 - создание спецификации источников данных и наполнение базы знаний.
- Верификация и оценка качества полученной базы знаний.

Все задачи трех этапов проекта МЕТЕОР выполнены.

Результаты работы по третьему этапу проекта за 2019 год нашли отражение в 6 научных публикациях (4 на английском языке), из них 3 статьи в рецензируемых научных журналах (1 в зарубежном журнале), и 3 докладов на международных научных конференциях. Публикации проиндексированы: в РИНЦ – 4, в Scopus – 4, в WoS – 1.

Всего по проекту – 25 публикаций (13 из них на английском языке). Из числа всех публикаций 6 статей в реферируемых журналах, 1 глава в монографии издательства Springer, и 18 докладов на международных и российских конференциях.

НАУЧНО-ПОПУЛЯРНОЕ ОПИСАНИЕ РЕЗУЛЬТАТОВ ПРОЕКТА

Представьте цифровое пространство, в котором можно найти и проанализировать информацию о любом объекте или человеке, и где данные связаны друг с другом и представлены в удобном для анализа виде. Например, представим виртуальный личный кабинет преподавателя. В нем находятся как личные данные о преподавателе, его доклады и статьи, проекты, конференции, читаемые курсы, аудиторная и внеаудиторная нагрузка, оценки студентов т.п. Рассматривая подобный кабинет в качестве отправной точки, можно узнать все о каждой публикации, о соавторах и цитируемых работах, далее о журнале, включая перечень опубликованных статей за все годы, затем об издательстве или о любом авторе и т.д. Если пойти по другой ветке, то можно получить информацию о дисциплинах, кто из студентов посещал каждую из дисциплин, информацию о каждом студенте и т.д.

Именно онтологии могут служить инструментом для создания подобных потенциально бесконечных связанных структур. Но онтологии представляют собой лишь безжизненный каркас, который необходимо наполнить данными, чтобы вдохнуть в него жизнь.

Причем эти данные уже существуют в различных базах данных, электронных таблицах и других информационных ресурсах. Проект МЕТЕОР предлагает новый метод наполнения онтологий из таких гетерогенных структурированных «родников». Основная идея и отличие этого метода от созданных ранее заключается в том, что информация об источниках данных фиксируется не на каком-то специализированном языке, а на языке онтологий с помощью предложенной онтологии ресурсных «родников» данных. Также метод фокусируется на решении проблемы наполнения онтологий из источников различного типа, в то время как его предшественники преимущественно были сосредоточены на включении в онтологию данных из источников одного определенного типа, например, реляционных баз данных.

Метод состоит из 6 шагов.

1. **Создание онтологии.** На этом этапе инженер по знаниям совместно с экспертами предметной области формулирует цель создания онтологии, и формируют начальную онтологию: выделяют ключевые концепты и описывают связи между ними, задают аксиомы и правила вывода новых знаний. При разработке начальной онтологии могут быть использованы как любые методологии, так и любые универсальные или расширенные редакторы онтологий. При выборе инструмента для работы с онтологиями рекомендуется учитывать особенности организации и класс задач, для решения которых планируется использовать онтологии [Беглер, Костоусов, 2018].
2. **Определение источников данных.** На данный момент в организациях любого типа собраны огромные массивы данных и цель этого этапа — выявить существующие источники данных, необходимые для решения задач согласно цели онтологии. В качестве источников могут выступать базы

данных, электронные таблицы, XML-файлы и другие файлы проприетарного формата, в которых хранятся структурированные данные.

3. **Формирование спецификации источников данных.** На этом этапе описываются свойства выявленных на предыдущем шаге источников данных. В частности, указывается тип, путь, название и описывается структура данных. Особенностью предлагаемого метода является отсутствие специализированного языка описания источников данных. Вместо этого предлагается специфицировать источники данных с помощью легковесной онтологии, которая задает структуру для всех свойств источников данных.
4. **Извлечение данных.** Данные из разнородных источников извлекаются, очищаются и объединяются в единую структуру. На этом этапе также выполняется проверка согласованности.
5. **Отображение онтологий.** Структура, созданная на предыдущем шаге, отличается от структуры онтологии предметной области, так как базируется на структуре источников данных. Для правильного импорта данных необходимо описать правила отображения онтологий. Описание правил осуществляется инженером по знаниям совместно с экспертом по предметной области с помощью встроенных в онтологию средств для аннотирования.
6. **Наполнение онтологии.** Используя созданную на предыдущем шаге спецификацию отображения онтологий, разработанная на первом шаге онтология наполняется данными из источников, выбранных на втором шаге.

В отличие от существующих методов, предполагающих однократное наполнение онтологии, предлагаемый метод можно использовать для односторонней синхронизации данных (причем это не требует повторения шагов 1-3 и 5, а шаги 4 и 6 выполняются автоматически). Также этот метод делает возможным одновременное наполнение онтологии данными из источников различных типов (с учетом ее семантики), при этом не требует изучения синтаксиса специфического языка описания отображений и/или разработки инструментария для облегчения создания описания отображения.

Разработанная методология МЕТЕОР может быть положена в основу систем поддержки принятия решений, основанных на знаниях. Она направлена на осуществление однотипных операций или на решение повторяющихся задач, требующих знания правил и процедур, а также сопоставления разнородной информации. Например, в медицинской онтологии могут быть заложены знания о необходимых анализах пациентов с различными диагнозами. Данные о пациентах и об анализах могут храниться в различных базах данных. До визита к врачу система, использующая методику МЕТЕОР, может «втянуть» в онтологию данные о пациенте из имеющихся источников и порекомендовать, какие анализы необходимо сдать заранее (или пройти обследования), чтобы визит к специалисту был более эффективным.

Перспективным приложением методологии являются также рекомендуемые системы, получившим широкое распространение в цифровом маркетинге и коммерции. Сейчас используются различные подходы к составлению списка рекомендаций: коллаборативные рекомендации, рекомендации на основе содержимого, рекомендации на основе знаний. Последний подход может быть реализован, в частности, с помощью

методологии МЕТЕОР. В онтологию закладываются знания о критериях выбора, совместимости объектов и т.п. Характеристики имеющихся объектов могут храниться в базе данных или отдельных файлах. Информация о предпочтениях пользователя содержится в запросе. По запросу данные об объектах «втягиваются» в онтологию, анализируются в соответствии с заложенными правилами и на основе проведенного анализа формируется рекомендация для конкретного пользователя.

ЗАЯВЛЕННЫЕ ЦЕЛИ ПРОЕКТА

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 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 связанных данных [<https://www.w3.org/OWL/>; <https://www.w3.org/standards/semanticweb/data>].

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.

ПОЛУЧЕННЫЕ В ПРОЕКТЕ ВАЖНЕЙШИЕ РЕЗУЛЬТАТЫ

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 первого этапа был проведен анализ существующих подходов к формированию корпоративных онтологических баз знаний и к интеграции имеющихся данных в созданную онтологию. Были получены следующие научные результаты (подробно описаны в отчете за 2017 год):

1-А) На основании анализа состояния мировых исследований в области онтологического инжиниринга были выявлены основные научные тренды в области баз знаний онтологического типа, включающие:

- существующие программные средства и технологии для создания баз знаний онтологического типа;
- особенности современных подходов к интеграции гетерогенных данных в онтологии и таксономии;
- анализ возможностей и требований, предъявляемых к языкам описания отображения (mapping description languages), а также их сравнительный анализ;
- проблемы, возникающие при интеграции данных на различных этапах жизненного цикла онтологии.

1-В) Была предложена методология МЕТЕОР, включающая:

- выбор инструментов для работы с корпоративными онтологиями;
- анализ и сравнение основных методов формирования и наполнения онтологий;
- обоснование методологии формирования онтологий на основе интеграции данных из гетерогенных источников с учетом когнитивных аспектов.

В 2018 году выполнялся **второй этап** проекта по созданию новой технологии формирования баз знаний на основе онтологий, основанной на современных подходах к интеграции, учитывающих семантику предметной области (Domain Semantics-Driven Mappings). Этап был посвящен разработке новой технологии, позволяющей компаниям эффективно наполнять базы знаний, интегрируя в них ранее накопленную информацию и данные. Разработанный новый подход учитывает особенности жизненного цикла онтологий, в частности, проблему устаревания данных. Также была предложена типовая архитектура системы.

На втором этапе были получены следующие научные результаты (подробнее описаны в отчете за 2018 год):

2-А) Предложена новая технология МЕТЕОР для формирования онтологий и интеграции структурированных данных из гетерогенных источников с учетом когнитивных аспектов, включающая:

- формальное описание методологии формирования и наполнения онтологий;
- создание онтологии источников данных различных типов (реляционная база данных, XML-документ, электронная таблица);
- разработка рекомендаций по описанию источников данных с использованием созданных онтологий.

Суть метода заключается в спецификации источников данных с помощью разработанной онтологии источников данных различного типа (мета-шаблона), а также описания отображения структуры данных в онтологию

предметной области в онтологической форме. Затем созданная спецификация и описание отображения используются для наполнения онтологии и последующей синхронизации данных, помогая решить проблему устаревания данных. Метод позволяет осуществлять одновременное наполнение онтологии данными из источников различных типов, при этом не требует изучения синтаксиса специфического языка описания отображений и/или разработки инструментария для облегчения создания описания отображения.

2-В) Также на втором этапе была создана архитектура и пилотный прототип системы на языке Python.

Третий заключительный этап (2019) был посвящен разработке, апробации и верификации прототипной системы МЕТЕОР-демо, реализующей созданные методологию и технологии на примере создания базы знаний учебного процесса ВУЗа и других предметных областей (медицина и сборочное производство).

На этом этапе были получены следующие основные результаты (в соответствии с планом из отчета за 2018):

3-А) Анализ административных задач в рамках учебных процессов и выбор задачи для кастомизации методологии создания базы знаний и ее архитектуры.

Был осуществлен анализ административных задач, возникающих в рамках управления и обеспечения учебных процессов. Эти задачи можно разделить на непосредственно связанные с функционированием учебного процесса (иными словами, «основные виды деятельности»), и задачи обеспечивающего (вспомогательного) характера (административно-хозяйственные, финансовые, кадровые и т.п.), являющиеся не специфическими по отношению к образовательной деятельности, а общими для деятельности любых организаций [Madzik et al., 2019; Beard, Humphrey, 2014].

Соответствующие основным видам деятельности задачи и предназначенные для их решения бизнес-процессы, в свою очередь, вслед за [Zhukova et al., 2014] можно подразделить на:

- учебные (в качестве объекта управления имеющие контингент студентов, проходящих обучение в образовательном учреждении);
- учебно-методические (в свою очередь, имеющие в качестве объекта управления учебно-методические материалы (в частности, образовательные стандарты, учебные планы и рабочие программы дисциплин) и профессорско-преподавательский состав).

Подобная классификация позволяет рассмотреть управление образовательными процессами с точки зрения различных заинтересованных сторон: обучающихся как непосредственных получателей образовательных услуг, и иных групп, оказывающих воздействие на формирование образовательных процессов.

Классификация проблем, в свою очередь, может основываться на классификации барьеров к обмену знаниями, предложенной в [Blagov et al., 2017], выделяющей барьеры информационно-технологического и организационно-управленческого характера. Такое разделение может в дальнейшем декомпозироваться на более частные проблемы. Связь данной классификации с исследованиями в области обмена знаниями достаточно актуальна в связи с развитием на базе ведущих учреждений высшего образования научно-образовательных центров (НОЦ), объединяющих образовательные учреждения с

представителями бизнеса и органами государственной власти. Представляется, что объединение исследований в области онтологического моделирования процессов управления и обеспечения образовательных процессов и поддержки обмена знаниями между различными стейкхолдерами данных процессов является потенциально плодотворным направлением развития исследований.

Таким образом, предлагаемая классификация задач, возникающих в рамках управления и обеспечения образовательного процесса, основывается (на верхнем уровне, с возможностью глубокой декомпозиции) на объекте управляющего воздействия; предлагаемая классификация проблем, осложняющих выполнение данных задач, основывается (опять же, с возможностью глубокой декомпозиции) на природе данных проблем. Сочетание данных двух классификаций может использоваться для построения схем матричного вида, классифицирующих задачи и проблемы при решении данных задач по двум указанным категориям.

Основные категории выявленных проблем следующие:

- проблемы сбора, хранения, обработки и получения информации;
- проблемы функциональности существующих информационных систем и технологий;
- проблемы, связанные с регламентами бизнес-процессов;
- проблемы коммуникации, распространения и доставки информации.

К критическим задачам или бизнес-процессам учебно-методического характера относятся:

- подготовка и реализация компетентностно-ориентированных учебных планов (КОУП);
- разработка и утверждение рабочих программ дисциплин (РПД);
- планирование нагрузки ППС;
- формирование блоков элективных дисциплин и организация записи на них студентов.

В качестве конкретной задачи для кастомизации и апробации методологии МЕТЕОР была выбрана задача формирования индивидуальных образовательных траекторий обучающихся, так как она тесно перекликается со всеми перечисленными выше задачами, при этом носит «человеко-центрированный» характер, то есть ориентирована на студента.

Решение данной задачи потребовало провести глубокий анализ компетенций, индивидуальных предпочтений и других компонентов «цифрового следа» обучающегося. Подобная необходимость анализа большого количества разнородных данных создаёт для данной задачи особую актуальность использования баз данных на основе онтологического моделирования; методологические основы работы коллектива авторов настоящего исследования над данной темой представлены в работе [Благов, Лещёва, Щербан, 2018].

Актуальность данной задачи обусловлена также и тенденциями современного развития образовательной сферы. Так, концепция Национальной Технологической Инициативы (НТИ) предполагает в качестве одного из основных принципов подготовки кадров для научно-технического прорыва на формирующихся рынках именно формирование индивидуальных треков непрерывного обучения на основе обработки с помощью технологий искусственного интеллекта (ИИ) «цифрового следа» обучающегося. Методика, предлагаемая в работе [Благов, Лещёва, Щербан, 2018] ориентирована в

первую очередь на создание индивидуальной образовательной траектории в рамках обучения на конкретной основной образовательной программе (ООП) уровня образования «бакалавриат»; дальнейшее развитие методики в направлении возможности применения для выстраивания индивидуальных образовательных траекторий непрерывного обучения на протяжении жизни (возможно, с учётом не только «цифрового следа» обучающегося, но и индивидуальных психологических черт либо же «когнитивных стилей» обучающегося [Gavrilova et al., 2013; Althuizen, 2018]) представляется весьма перспективным направлением развития исследований в данной области.

3-В) Апробация новой технологии на предметных областях.

і) Предметная область «Образование»

Для решения выбранной задачи и апробации разработанного в рамках проекта подхода был создан прототип базы знаний, концептуальная схема которой частично представлена на рисунке 1.

Источники документации для наполнения онтологии включают в себя:

- профессиональные стандарты (как источник трудовых функций);
- федеральные государственные образовательные стандарты (ФГОС) по соответствующим направлениям подготовки (как источник формулировок компетенций универсального, общепрофессионального, профессионального и профильного характера);
- компетентностно-ориентированные учебные планы (КОУП) как источник перечня дисциплин по различным направлениям и профилям обучения;
- рабочие программы дисциплин (РПД) как источник перечня характеризующих каждую конкретную дисциплину: а) компетенций; б) результатов обучения (знаний, умений и навыков).

Использование документов указанных категорий (перечисленных в порядке от общего к частному) позволяет выстроить полноценную онтологию конкретной образовательной программы, которая может использоваться следующими категориями пользователей для решения следующих категорий задач:

- руководством программы для наглядного представления логики построения программы и взаимосвязей между её элементами, что может позволить быстро принимать логически обоснованные решения по модификации программы как с точки зрения набора входящих в неё элементов (напр., добавление или удаление дисциплин из/в КОУП), так и с точки зрения взаимосвязи между этими элементами (напр., изменение очередности преподавания конкретных дисциплин) для более успешного освоения обучающимися конкретных компетенций и результатов обучения;
- обучающимися для формирования индивидуальных траекторий обучения с преимущественной ориентацией на освоение конкретных результатов обучения;
- работодателями для оценки конкретных компетенций, знаний, навыков и умения выпускников образовательной программы, освоивших данную программу с использованием конкретной индивидуальной образовательной траектории.

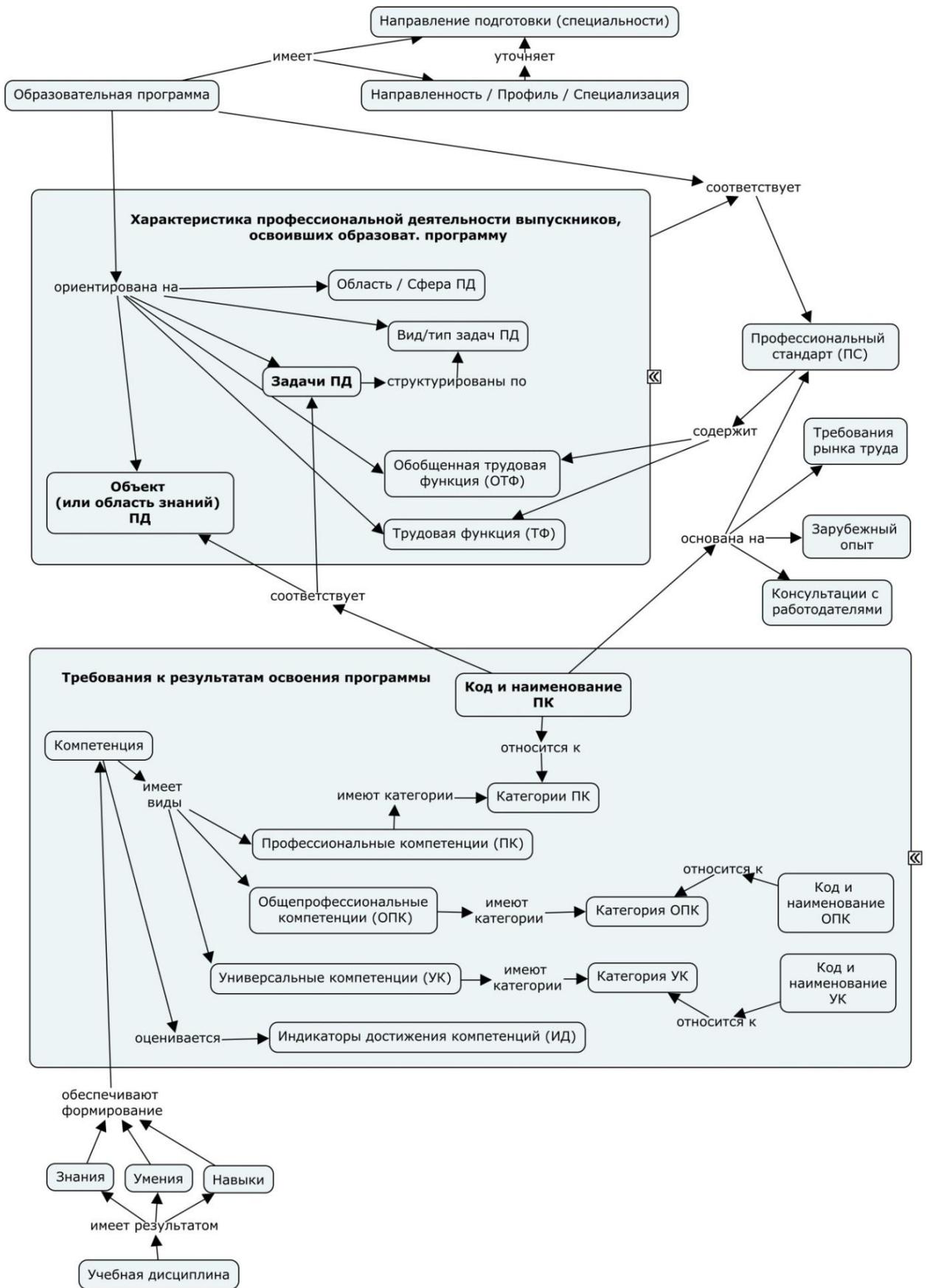


Рис. 1. Концептуальная модель формирования компетенций

ii) Апробация на предметной области «Медицина»

Также была осуществлена апробация методологии МЕТЕОР для формирования и наполнения онтологии диагностики орфанных заболеваний. Онтологический подход для диагностики орфанных заболеваний было предложено использовать в работе [Кобринский и др., 2019]. Диагностирование осуществляется на основе шкал модальности признаков с их коэффициентами и формул для комплексной оценки совокупности факторов уверенности (выраженности и манифестации).

Для реализации предложенного подхода была создана онтология *ontomorf* (на основании материалов Б.А. Кобринского [Кобринский, 2018]), основные элементы которой представлены на рисунке 2.

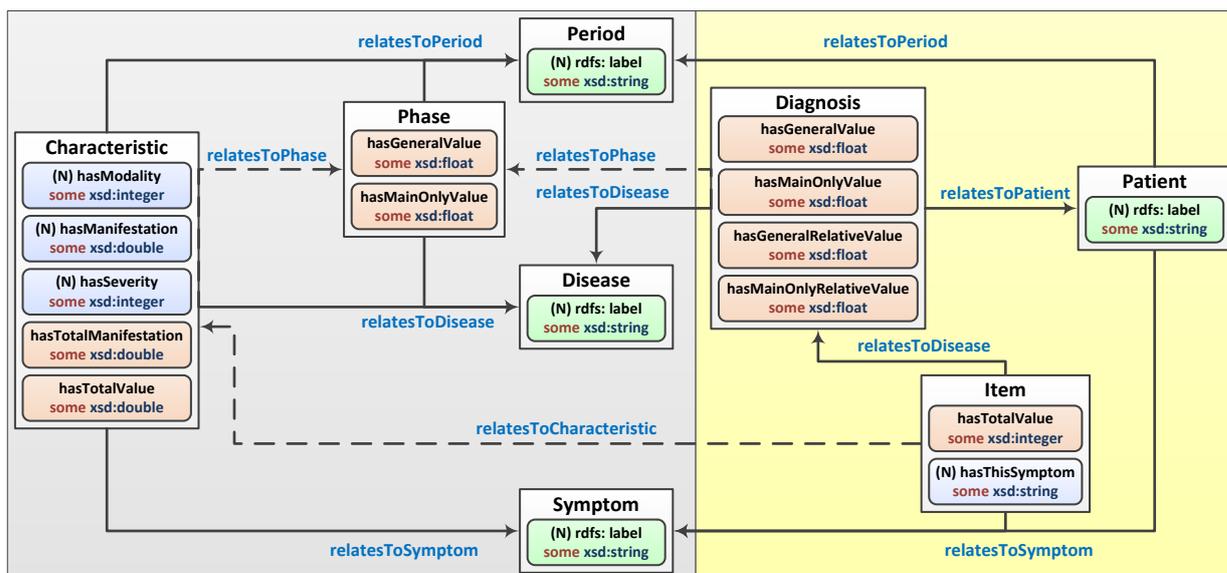


Рис. 2. Онтология диагностики орфанных заболеваний

В левой части — классы, в которых содержится «постоянная» информация базы знаний. Предполагается, что в процессе работы базы знаний эта информация не меняется или меняется очень редко. Правая часть содержит «переменную» информацию о пациентах.

Разработанный метод МЕТЕОР при наполнении онтологии *ontomorf* использовался несколько раз:

- 1) на этапе создания «постоянной» части базы знаний (однократно);
- 2) на этапе оценки для внесения информации о пациентах.

Данные для наполнения базы знаний содержались в Excel-файлах. Для вычислений и логического вывода на базе был сформулирован набор правил (24 правила). Пример правила следует ниже:

```
Diagnosis(?dg), Disease(?di), relatesToDisease(?dg, ?di), Phase(?ph),  
relatesToDisease(?ph, ?di), Period(?pe), relatesToPeriod(?ph, ?pe),  
Patient(?pa), relatesToPatient(?dg, ?pa), relatesToPeriod(?pa, ?pe) ->  
relatesToPhase(?dg, ?ph)
```

С помощью подобных правил осуществлялись вычисления значения свойств (например, суммарной манифестации, степени уверенности в диагнозе и т.п.) и устанавливались связи между объектами, показанные пунктирными линиями на рисунке 2.

Для оценки результатов диагностики на созданной базе, были взяты реальные истории болезни пациентов из отечественных и зарубежных открытых источников с верифицированными диагнозами. Диагностика с помощью сформированной базы знаний продемонстрировала (по оценке независимых экспертов) вполне приемлемые результаты для до-лабораторного этапа идентификации заболевания.

iii) Апробация на предметной области «Производство»

Задача создания и наполнения онтологического графа знаний (см. рис. 3) для базы знаний автомобилестроительной компании возникла в связи с промышленной необходимостью. Метод МЕТЕОР был применён для формирования и наполнения пилотной БЗ из примерно 50 экземпляров сборочных единиц и выделенных в процессе создания около 400 свойств. Элементы базы знаний представляют собой сборочные единицы (СЕ) различных моделей автомобилей, в основном электрогрузовиков, производимых компанией. Данные для наполнения базы знаний содержались в онлайн справочниках, каталогах, ГОСТах, рабочих Excel-файлах и базах данных. Созданный граф знаний служит основой для новой корпоративной системы управления знаниями, в которой пользователи (инженер, менеджер по закупкам, экономист и т. д.) могут найти СЕ, в которой они нуждаются, а также всю соответствующую информацию. Создание подобной системы имеет принципиальное значение для компании, поскольку выпускаемые автомобили преимущественно собираются из СЕ, доступных на рынке, поиск информации о которых является одной из наиболее затратных по времени задач, решаемой инженерами в процессе проектирования. Предполагается, что создаваемая система будет интегрировать как информацию о СЕ, уже используемых в компании, так и о тех, которые представлены на рынке в каталогах поставщиков. В будущем база знаний может использоваться в роботизированном производстве, которое также развивается в компании.

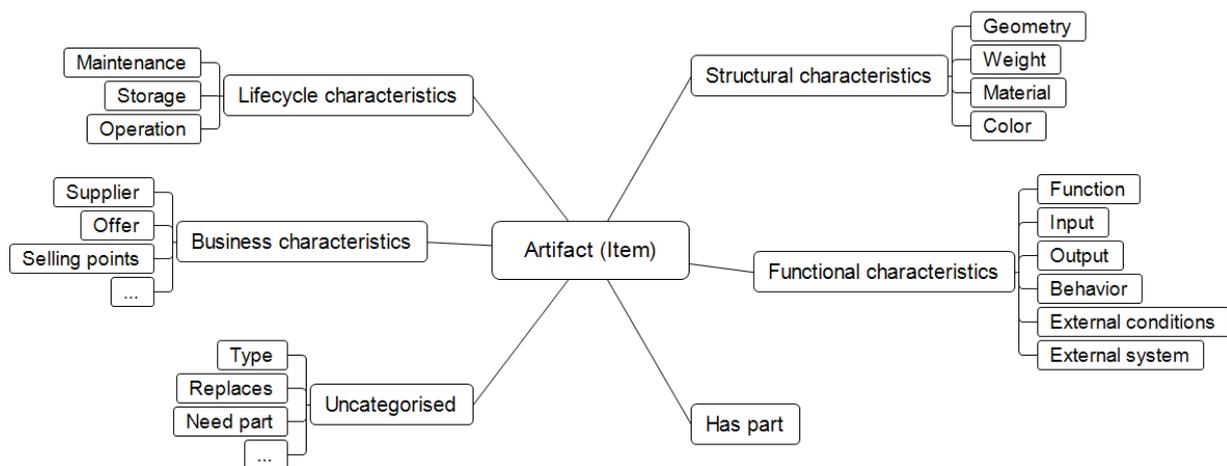


Рис. 3. Макет Онтологии сборочных единиц (далее был доработан с использованием Protege)

3-С) Верификация и оценка качества полученной базы знаний.

Три разработанные онтологии перед их непосредственным наполнением и использованием в бизнес-процессах были подвергнуты процедурам верификации и

оценки. Оценка (evaluation) онтологии включает в себя два процесса: собственно верификацию (verification) и валидацию (validation).

Термин «оценка онтологии» подразумевает вынесение суждений о технической реализации онтологии, её программном обеспечении и документации и соответствии требованиям (например, способности отвечать на компетентностные вопросы).

Верификация онтологии относится к проверке технической реализации онтологии, программного обеспечения и документации, в то время, как валидация онтологии должна гарантировать то, что онтология, программное обеспечение и документация действительно представляют собой ту систему, которую они должны представлять. Кроме этого, проверка онтологии включает её экспертизу (assessment) – возможность её использования для удовлетворения потребностей пользователя [Gomez-Perez, 1995].

Большинство подходов к оценке онтологий попадают в одну из четырёх категорий [Brank et al., 2005]:

1. Сравнение онтологии с «золотым стандартом».
2. Использование онтологии в приложении и оценка результата его работы.
3. Сравнение онтологии с источником данных (например, документами из доменной области, которую должна описать онтология).
4. Экспертная оценка людьми на предмет соответствия predetermined требованиям и стандартам.

Оценка разработанных онтологий и их наполнения проводилась экспертами из предметных областей совместно с коллективом участников проекта. Процесс носил итеративный неформализованный характер и с доработками и исправлениями занял приблизительно 20% от общего времени. Подробнее особенности коллективной разработки описаны в работе [Кудрявцев, Беглер, Гаврилова, Лещева и др., 2019].

Таким образом все задачи проекта выполнены, а результаты отражены в 25 публикациях.

СОПОСТАВЛЕНИЕ ПОЛУЧЕННЫХ РЕЗУЛЬТАТОВ С МИРОВЫМ УРОВНЕМ

Онтологии вот уже более 20 лет являются одной из наиболее перспективных моделей представления знаний [Петрова, Тузовский, 2015; Боргест, 2018, Юрин, Дородных, Коршунов, 2018]. Однако до настоящего момента узким местом исследований является автоматизированное расширение (наполнение и обогащение) существующей онтологии, как правило, на основе информации, извлеченной из предметно-ориентированного контента. Под наполнением онтологии (ontology population) понимают добавление в онтологию индивидов с их свойствами, а под обогащением (ontology enrichment) — добавление новых отношений и аксиом, использующих эти отношения. Проблематика наполнения и обогащения онтологий является достаточно популярной областью исследований. На настоящий момент это одно из наиболее как теоретически и методологически глубоких, так и непосредственно практически ориентированных направлений управления знаниями. Однако, как указано выше, существующие методологии отличаются целым рядом недостатков, придающих технологии проекта МЕТЕОР существенную степень уникальности. Большинство существующих подходов направлены на популяцию онтологии данными из определённого типа источников.

Так, начиная с 1998 года, когда появилась первая версия XML, было реализовано множество разработок и исследований, связанных с преобразованием XML-файлов в формальные онтологии, представленные в форматах OWL и RDF [Erdmann and Studer, 2001; Thuy et al., 2008; Aussenac-Gilles, Kamel, 2009; Bakkas, Jakjoud, Bahaj, 2014]. Согласно [Hacherouf, Bahloul, Cruz, 2015] существующие подходы можно разделить на две основные группы. Первая группа, это подходы, ориентированные на экземпляры (instance approaches), которые позволяют работать с произвольными XML-документами без наличия соответствующей XML-схемы [JXML2OWL, XMLMaster]. Ко второй группе относятся валидационные подходы (validation approaches), основанные на обработке XML-схем, используемых для валидации соответствующих XML-документов.

Данные табличного типа (tabular form) широко распространены и используются рядом приложений: текстовыми процессорами (text processor), электронными таблицами (spreadsheet), программами для подготовки презентаций (presentation program) и языками разметки (markup language). Большинство существующих решений для интеграции данных рассматривают электронные таблицы и языки разметки, в частности HTML [Ozturk, 2019].

Таблицы включают в себя как данные (эксплицитно), так и их схему (имплицитно, в структуре строк и столбцов). Таким образом, из таблиц можно извлекать как данные (ABox) для наполнения онтологии, так и непосредственно структуру онтологии (TBox) для её обогащения, а также свойства (RBox) [Compton et al., 2015]. Существующие решения охватывают как извлечение схемы источника данных (например, [Tijerino et al. 2005; An, Park, 2018] для HTML-таблиц), так и извлечение данных в ранее созданную схему. Принципиальная разница в интеграции данных этих двух подходов в том, откуда берётся структура получившейся базы данных – в первом случае это структура источника данных (direct mapping), во втором она создаётся заранее (domain semantic-driven mapping).

Многочисленные исследования, посвященные наполнению онтологий из реляционных баз данных (например, [Cullot, Ghawi, Yetongnon, 2007; De Laborda, Conrad, 2006; Svihla, Jelinek, 2004]), привели к созданию рабочей группой W3C рекомендаций A Direct Mapping of Relational Data to RDF [<https://www.w3.org/TR/rdb-direct-mapping/>] (Прямое отображение реляционных данных в RDF) и R2RML: RDB to RDF Mapping Language [<https://www.w3.org/TR/r2rml/>] (R2RML: Язык отображения RDB в RDF). Первая рекомендация регламентирует прямое отображение реляционных баз данных в RDF. Второй 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>) и DB2Triples (<https://github.com/antidot/db2triples>).

При наполнении существующей онтологии данными из структурированных источников (domain semantic-driven mapping) правила отображения задаются явным образом с помощью языков описания отображения (mapping description languages) [Calvanese, et al. 2018.]. Кроме уже упомянутого выше R2RML, известен целый ряд подобных языков: DR2 MAP [Bizer, 2003], R2O [Barrasa, Corcho, Gómez-Pérez, 2004], DR2Q [Cyganiak et al., 2012] для интеграции с базами данных, для консолидации данных

из XML [Bohring, Auer, 2015; Rodrigues, Rosa, Cardoso, 2006; Anicic, Ivezic, Marjanovic, 2007], для интеграции данных из электронных таблиц [Tijerino et al., 2005].

Для решения проблемы интеграции данных в онтологию предлагались также подходы, не использующие языки отображения. Например, RDBToOnto [Cerbah, 2008] анализирует избыточность данных, чтобы выявить шаблоны категоризации. А в работе [Hu, Qi, 2007] применяются методы интеллектуального анализа данных для автоматического обнаружения отображений между базой данных и существующей онтологией.

Интеграция данных в онтологию может осуществляться как методом отображения по требованию (mapping on-demand), так и включением данных (консолидация) в базу знаний в качестве значений свойств объектов (data materialization). Второй метод используется, в частности, при построении хранилищ данных (data warehouse), но при его использовании необходимо учитывать проблему устаревания данных.

С технической точки зрения разрабатываемая технология является расширением и дополнением существующих методов консолидации данных и наполнения онтологий, учитывающих семантику предметной области. В частности, в отличие от методов, рассматриваемых и предлагаемых в [Бова, 2015; Chungooga et al., 2013; Sitharamulu, Babu, 2016; Paulheim, 2017], технология проекта МЕТЕОР позволяет производить не только однократное наполнение онтологии, но и синхронизацию данных и наполнение онтологии данными различных форматов и источников.

В результате реализации проекта МЕТЕОР создана новая методология, позволяющая гибко настраивать интеграцию данных из источников различного типа в ранее созданную онтологию; поддерживающая модульность и учитывающая проблему устаревания данных. Эта методология позволяет проектировать и поддерживать современные корпоративные базы знаний для эффективной цифровой трансформации и информационного менеджмента.

К более специфичным чертам разрабатываемой методологии (впрочем, непосредственно связанным с её общей идеологией), отличающим её от других подходов, можно отнести то, что при формировании спецификации источников данных для наполнения онтологий не предполагается формирование специализированного языка описания отображений, вместо которого используется мета-шаблон, представляющий собой легковесную онтологию; таким образом, разрабатываемая методология приобретает дополнительную гибкость.

Соответственно, можно заключить, что разрабатываемая в рамках проекта МЕТЕОР методология, находясь в русле ключевых и наиболее динамично развивающихся направлений исследований в области управления знаниями в целом и построения и поддержки онтологий в частности, отличается рядом черт общей архитектуры и конкретных решений. Эти черты позволяют максимально гибко интегрировать в онтологии данные различных типов с учётом проблем устаревания данных и когнитивных аспектов дизайна и генерации онтологий. Данная гибкость придаёт результатам проекта МЕТЕОР, в частности, его второго этапа, потенциально высокую степень практической применимости для управления знаниями организаций различных отраслей и бизнес-моделей. Представляется, что помимо образовательной отрасли, особенную актуальность проблематика объединения в рамках единой онтологии разнородных данных имеет в российском контексте в отношении производственных отраслей – вследствие кадрового

кризиса в период 1990-х гг. эти области испытывают затруднения с межвременным сохранением и передачей технологических и организационных знаний [Бодрова, Голованова, 2017; Балацкий, Юревич, 2018]. Таким образом, результаты работы, помимо теоретической и методологической, имеют и безусловную практическую значимость для различных отраслевых контекстов.

Представляется, что описанные отличия предлагаемой методологии от российских и зарубежных аналогов делают ее перспективной с точки зрения внедрения в практику деятельности организаций различных сфер деятельности, в частности, к примеру, образовательных учреждений [Благов, Лещёва, Щербан, 2018], эффективность и результативность деятельности которых в особенной степени зависит от способности максимально логичным и непротиворечивым образом структурировать используемые в работе данные, информацию и знания различного характера.

Ценность полученных результатов подтверждается и успешным представлением данных результатов на различных как российских, так и международных научных и научно-практических конференциях.

Новизну разработанных в рамках проекта МЕТЕОР методов и полученных результатов можно подразделить на новизну по отношению к мировому контексту исследований в области онтологического моделирования корпоративных знаний в целом и по отношению к российскому контексту подобных исследований.

Относительно новизны в мировом контексте следует отметить ориентацию разрабатываемой методологии на преодоление таких недостатков существующих подходов к наполнению и обогащению онтологий как [Petasis et al., 2011; Nederstigt et al., 2014; Otero-Cerdeira, L. et al., 2015]: узкая направленность на решение специфических задач конкретных проектов, обеспечение возможности интеграции данных только из однотипных источников, отсутствие модульности и расширяемости и недостаток удобного инструментария. Предлагаемая методология позволяет использовать для интеграции данные различных типов (в частности, реляционные базы данных, XML-документы, электронные таблицы MS Excel) и таким образом создавать и модифицировать, в т.ч., расширять, онтологии различных предметных областей.

Новизна разрабатываемой методологии в российском контексте заключается в самом фокусе на интеграцию разнородных данных в онтологии предметной области с учётом когнитивных аспектов дизайна и генерации онтологий. Большинство российских работ, посвящённых интеграции данных с использованием онтологий (напр., работы коллективов под руководством В.Ф. Тузовского, М.Р. Когаловского. Дородных Н.О., и др.), используют онтологии в качестве инструмента, но не предполагают моделей и методов создания новых, либо модификации существующих онтологий в качестве результата работы. Таким образом, предлагаемая в рамках проекта МЕТЕОР методология в контексте российских исследований схожей тематики обладает высокой степенью новизны и актуальности как с теоретической и методологической точки зрения, так и с точки зрения практической реализации разработанного пилотного прототипа, позволяющего решать задачи, до этого не решавшиеся программными средствами российских разработчиков.

МЕТОДЫ И ПОДХОДЫ, ИСПОЛЬЗОВАННЫЕ ПРИ РЕАЛИЗАЦИИ ПРОЕКТА

Разработка нового метода наполнения онтологий и интеграции структурированных данных из гетерогенных источников была основана на системном подходе и на существующих технологиях интеграции, учитывающих семантику предметной области (Domain Semantics-Driven Mappings), использующих языки описания отображений (mapping description languages).

Проект Метеор в целом выполнялся в рамках методологии «проектирующего исследования» (design science research), в соответствии с которой исследователь решает актуальные проблемы путем создания инновационных артефактов, пригодных для повторного использования и вносящих вклад в научное знание [Hevner, Chatterjee, 2010].

«Проектирующее исследование» в соответствии с [Österle et al, 2011] должно следовать 4-м основным принципам:

- **Обобщенность** – создаваемый артефакт применим к классу проблем;
- **Оригинальность** – создаваемый артефакт должен расширять существующие знания и предмете;
- **Обоснованность** – создаваемый артефакт должен быть подкреплён понятными аргументами и предоставлять возможность валидации;
- **Полезность** – создаваемый артефакт должен приносить пользу – либо в ближайшее время, либо в будущем – для соответствующих заинтересованных сторон.

В соответствии с [Peppers et al, 2007] «проектирующее исследование» включает 6 шагов, которым следовал авторский коллектив при разработке методологии МЕТЕОР:

1. Идентификация и обоснование проблемы (актуальность);
2. Идентификация целей, требований и ограничений;
3. Проектирование и разработка артефакта;
4. Демонстрация использования (апробация) созданного артефакта;
5. Оценка созданного артефакта с точки зрения его результативности, эффективности и т.п.;
6. Обнародование полученных результатов.

В целом задачи проекта были направлены на широкое аналитическое исследование и разработку новых алгоритмов в области онтологического инжиниринга. Для проведения этих работ использовались следующие методы:

- контекстный анализ;
- системный анализ;
- критический структурный обзор литературы, технологий, платформ и сервисов (50 источников, список приложен в файле pdf);
- структурный дизайн и концептуализация понятий с использованием интеллект-карт;
- инструментарий онтологического инжиниринга, в частности, методология разработки онтологий NeOn;
- инструменты Google Analytics;
- генерализация, категоризация и синтез понятий для построения таксономий.

Разработка нового метода наполнения онтологий и интеграции структурированных данных из гетерогенных источников была основана на системном подходе и на существующих технологиях интеграции, учитывающих семантику предметной области (Domain Semantics-Driven Mappings), использующих языки описания отображений (mapping description languages).

В сочетании с «проектирующим исследованием» эти методы использованы впервые, а их синергетический эффект позволил получить новые оригинальные результаты, опубликованные в 25 научных работах.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК ВСЕХ ПУБЛИКАЦИЙ ПО ПРОЕКТУ

Статьи

1. Гаврилова Т., Кудрявцев Д., Кузнецова А. Выбор инструментов управления знаниями с учетом специфики предметной области // *Инновации*. – 2019. – №8(250). – с. 44-52. (РИНЦ, список ВАК)
2. Кудрявцев Д.В., Беглер А.М., Гаврилова Т.А., Лещева И.А., Кубельский М.В., Тушканова О.Н. Метод визуальной коллективной разработки онтологического графа знаний // *Искусственный интеллект и принятие решений*. – 2019. – №1 – с. 27-38. DOI: 10.14357/20718594190103 (РИНЦ, список ВАК, WoS)
3. Gavrilova T., Kudryavtsev D., Grinberg E. Aesthetic Knowledge Diagrams: Bridging Understanding and Communication // In: Handzic M., Carlucci D. (eds) *Knowledge Management, Arts, and Humanities. Knowledge Management and Organizational Learning*, Vol. 7, Springer, Cham. - 2019. – pp. 97–117. DOI: 10.1007/978-3-030-10922-6_6 (РИНЦ, Scopus)
4. Благов Е.Ю., Лещёва И.А., Щербан С.А. Онтологический подход в практике образовательной деятельности: формирование траекторий индивидуального профессионального развития студентов // *Открытое образование*. – 2018. – №5. – с. 26-39. (РИНЦ, список ВАК)
5. Гринберг Э.Я., Плешкова А.Ю. К вопросу о подготовке специалистов по управлению знаниями // *Открытое образование*. – 2018. – №22(2). – с.14-26. DOI:10.21686/1818-4243-2018-2-14-26 (РИНЦ, список ВАК)
6. Gavrilova T., Kubelskiy M., Kudryavtsev D., Grinberg E. Modeling methods for strategy formulation in a turbulent environment // *Strategic Change (Wiley publishing)* - 2018. - Vol. 27, No. 4. - pp. 369-377. DOI: 10.1002/jsc.2209 (РИНЦ, Scopus, WoS)
7. Гаврилова Т.А., Онуфриев В. А. Анализ ошибок студентов при визуальном структурировании знаний // *Компьютерные инструменты в образовании*. – 2016. – № 6. – с. 42-54. (РИНЦ, список ВАК)

Тезисы и доклады

8. Gavrilova T., Grinberg E. Visual Ontology Sketching for Preliminary Knowledge Base Design // In: Bi Y., Bhatia R., Kapoor S. (eds) *Intelligent Systems and Applications. IntelliSys 2019. Advances in Intelligent Systems and Computing*, vol 1037. Springer, Cham, 2020. – pp. 566-576.
9. Gavrilova T., Kokoulina L. Using Ontology Engineering to Design Artificial Intelligence Course // In: *Smart Education and e-Learning SEEL, Smart Innovation, Systems and*

- Technologies series 144, V. L. Uskov et al., (eds), Springer, 2019. – pp. 201-207.
https://doi.org/10.1007/978-981-13-8260-4_19
10. Gavrilova T., Kuznetsova A. Developing a Big Picture View on Research through Visualization Techniques // Proceeding of 14th International Forum on Knowledge Asset Dynamics IFKAD 2019, Matera, 2019. – pp. 179-184. https://www.ifkad.org/wp/wp-content/uploads/documents/ebooks/IFKAD2019_PROCEEDINGS_eBOOK.pdf
 11. Кудрявцев Д.В., Беглер А.М. Системы организации знаний: виды и примеры использования // Сборник научных трудов XXI Российской научной конференции "Инжиниринг предприятий и управление знаниями" (ИП&УЗ-2018), Москва, 26-28 апреля 2018. – с.184-190.
 12. Кудрявцев Д.В., Гаврилова Т.А., Лещева И.А., Беглер А., Кубельский М., Тушканова О. Методика групповой работы по визуальной разработке графа знаний // Труды национальной конф. по искусственному интеллекту КИИ-2018, Москва, 2018. - Т.1. - с.53-61.
 13. Лещева И.А. Практический подход к созданию и наполнению баз знаний онтологического типа: проблемы и вызовы // Сборник научных трудов XXI Российской конференции «Инжиниринг предприятий и управление знаниями», М., ФГБОУ ВО «РЭУ им. Г.В. Плеханова», 2018. – Т.1. - С.194-201.
 14. Kudryavtsev D., Gavrilova T., Kuznetsova A The Efficacy of Knowledge Management Tools across Domains // Book of abstracts, 13th International Forum on Knowledge Asset Dynamics IFKAD 2018, Delft, 2018. – p.141. Online proceedings URL: https://www.ifkad.org/wp/wp-content/uploads/documents/papers/y2018/135_IFKAD2018.pdf P. 1785-1794.
 15. Kudryavtsev D., Gavrilova T., Leshcheva I., Begler A., Tushkanova O. Mind Mapping and Spreadsheets in Collaborative Design of Knowledge Graphs // Joint Proceedings of the BIR 2018 Short Papers, Workshops and Doctoral Consortium co-located with 17th International Conference Perspectives in Business Informatics Research (BIR 2018). Stockholm, Sweden Vol. 2218. - 2018. - P.82-93. <http://ceur-ws.org/Vol-2218/paper8.pdf>
 16. Kudryavtsev D., Gavrilova T., Leshcheva I., Begler A., Kubelskiy M., Tushkanova O. A method for collaborative visual creation of a knowledge graph // Conference Proceedings "GSOM Emerging Markets Conference 2018", St.Petersburg, Russia, 2018. - pp. 98-102.
 17. Leshcheva I., Leshchev D. Ontology as mapping of material world // A. Sarygulov, V. Sergeev, L. Ungvári and W. Semmler (Eds.) SHS Web of Conf. 44 00055 (2018) Proceedings of IV International Scientific Conference “The Convergence of Digital and Physical Worlds: Technological, Economic and Social Challenges” (CC-TEESC2018), St. Petersburg, Russia, May 16-18, 2018. - P. 1-9. DOI: 10.1051/shsconf/20184400055 URL: https://www.shs-conferences.org/articles/shsconf/pdf/2018/05/shsconf_cc-tesc2018_00055.pdf
 18. 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", 29-30 июня 2017 г., Санкт-Петербург, 2017. – с. 226-234.
 19. 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". Моск. госуд. ун-т экономики, статистики и информатики. - М., 2017. - с. 68-73.
20. 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.
 21. 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. — p. 102. Proceeding online: pp. 840-851.
 22. 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, 2017. — pp. 637-645.
 23. 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, Barcelona, 2017. — pp. 345-352.
 24. Leshcheva I., Blagov E., Pleshkova A. Towards a Method of Ontology Population from Heterogeneous Sources of Structured Data // Proc. of 30th Jubilee Neumann Colloquium, 24-25 November, Budapest, Hungary, 2017. – pp. 29-34.
 25. Onufriev V., Gavrilova T. Conceptual Modelling: Common Students' Mistakes In Visual Representation // Conference Proceeding of ICL2017 – Teaching and Learning in a Digital World, 27-29 September, Budapest, Hungary, 2017. — pp. 62-71.

УЧАСТИЕ В НАУЧНЫХ МЕРОПРИЯТИЯХ ПО ТЕМАТИКЕ ПРОЕКТА ЗА 2019 ГОД

1. 14th International Forum on Knowledge Asset Dynamics (IFKAD-2019) (14-ый Международный Форум по динамике знаниевых активов), 5-7 июня 2019 г., г. Матера, Италия, Университет Базиликаты (Università degli Studi della Basilicata).
2. International Conference Cyber-Physical Systems and Control (CPS&C'2019) (Международная Конференция по кибер-физическим системам и управлению), 10-12 июня 2019 г., г. Санкт-Петербург, Санкт-Петербургский политехнический университет Петра Великого.
3. 6th KES International Conference on Smart Education and E-Learning (SEEL-19) (6-я международная конференция по умному образованию и электронному обучению), 17-19 июня 2019 г., г. Сан-Джильян, Мальта, ассоциация «KES International».
4. Intelligent Systems Conference 2019 (IntelliSys-2019) (Конференция по интеллектуальным системам), 5-6 сентября 2019 г., г. Лондон, Великобритания, Ассоциация «Научные и информационные конференции» (SAI Conferences).
5. Semantics 2019 (Конференция Семантика), 10-11 сентября 2019, Карлсруэ, Германия, Leibniz Institute for Information Infrastructure (Институт информационной инфраструктуры Ассоциации Лейбница).
6. Научно-практическая конференция «Практические инструменты управления знаниями», 3 декабря, г. Москва, НИУ ВШЭ.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Балацкий Е. В., Юревич М. А. Моделирование возрастной структуры научных кадров // *Terra Economicus*. – 2018. – Т. 16. №3. – С. 60-76.
2. Беглер А. М., Костоусов, С. А. Критерии выбора инструмента по работе с корпоративными базами знаний // Сборник студенческих работ V Международной научно-практической конференции “Управленческие науки в современном мире” 2018 / г. Москва, (2018). – С. 18–23.
3. Благов Е.Ю., Лещева И.А., Щербан С.А. Онтологический подход в практике образовательной деятельности: формирование траекторий индивидуального профессионального развития студентов // *Открытое образование*. – 2018. – Т. 22. №5. – С. 26-39.
4. 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.
5. Бодрова Е. В., Голованова Н. Б. Модернизация высшей технической школы: исторический опыт и современность // *Российский технологический журнал*. – 2017. – Т. 5. № 6 (20). – С. 73-97.
6. Боргест Н. М. Онтологии проектирования от Витрувия до Виттиха // *Онтология проектирования*. – Т.8 №4 (30). - 2018. – С. 487-522.
7. Кобринский Б. А. Триединство факторов уверенности в задачах медицинской диагностики // *Искусственный интеллект и принятие решений*. – 2018. – №. 2. – С. 62-72.
8. Кобринский Б.А., Благодосклон Н.А., Демикова Н.С., Грибова В.В., Шалфеева Е.А., Петряева М.В. Возможности применения онтологического подхода к диагностике орфанных заболеваний// Семнадцатая Национальная конференция по искусственному интеллекту с международным участием. КИИ-2019 (21–25 октября 2019 г., г. Ульяновск, Россия). Сборник научных трудов. В 2 т. – Ульяновск: УлГТУ, 2019. – Т.2. – С. 227. ISBN 978-5-9795-1940-1.
9. Кудрявцев Д. В., Беглер А. М., Гаврилова Т. А., Лещева И. А., Кубельский М. В., Тушканова О.Н. Метод коллективной визуальной разработки онтологического графа знаний // *Искусственный интеллект и принятие решений*. – 2019. – №. 1. – С. 27-38.
10. Петрова Г.Г., Тузовский, А.Ф. Онтология FIBO и ее применение для решения прикладных задач финансовой сферы // *Знания-Онтологии-Теории (ЗОНТ-2015)*. – 2015. – С. 96-99.
11. Юрин А.Ю., Дородных Н.О., Коршунов С.А. Средства поддержки моделирования логических правил в нотации RVM // *Программные продукты и системы*. – 2018. – Т. 31, №4. – С. 667-672
12. Althuizen N. Using structural technology acceptance models to segment intended users of a new technology: Propositions and an empirical illustration // *Information Systems Journal*. – 2018. – Т. 28, № 5. – С. 879-904.
13. An J., Park Y.B. Methodology for automatic ontology generation using database schema information // *Mobile Information Systems*. – 2018. – Т.2018 – 13 С. ID 1359174. DOI 10.1155/2018/1359174.
14. Anicic N., Ivezic N., Marjanovic Z. Mapping XML schema to OWL // *Enterprise Interoperability*. – Springer, London, 2007. – С. 243-252.

15. Aussenac-Gilles N., Kamel M. *Ontology Learning by Analyzing XML Document Structure and Content // Knowledge Engineering and Ontology Development (KEOD)*. – Madeira, Portugal, 2009. – C. 159-165.
16. Bakkas J., Jakjoud W., Bahaj M. *Semantic mapping at the schema level of XML documents to ontologies // Next Generation Networks and Services (NGNS), 2014 Fifth International Conference on*. – IEEE, 2014. – C. 165-169.
17. 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. – C. 1069-1070.
18. Beard D. F., Humphrey R. L. *Alignment of University Information Technology Resources With the Malcolm Baldrige Results Criteria for Performance Excellence in Education: A Balanced Scorecard Approach // Journal of Education for Business*. – 2014. – T. 89, № 7. – C. 382-388.
19. Bizer C. *D2R MAP - a database to RDF mapping language // Proceedings of the 12th International World Wide Web Conference*. – Budapest, Hungary, 2003. – URL: <http://wwwconference.org/www2003/cdrom/papers/poster/p004/p4-bizer.html>
20. Blagov E., Pleshkova A., Soldatkin E., Koritckiy N. *Knowledge Sharing Barriers in the Educational Program Management Administrative Processes: A Case of a Bachelor Program in a Russian University // Electronic Journal of Knowledge Management*. – 2017. – T. 15, № 2. – C. 113-125.
21. Bohring H., Auer S. *Mapping XML to OWL ontologies // Marktplatz Internet: Von e-Learning bis e-Payment, 13. Leipziger Informatik-Tage (LIT 2005)*. – 2015. – C. 147-156
22. Brank J., Grobelnik M., Mladenic D. *A survey of ontology evaluation techniques // Proceedings of the conference on data mining and data warehouses (SiKDD 2005)*. – Citeseer Ljubljana, Slovenia, 2005. – C. 166-170.
23. Calvanese D., et al. *Efficient ontology-based data integration with canonical IRIs // European Semantic Web Conference*. – Springer, Cham, 2018. – C. 697-713.
24. Cerbah F. *Learning highly structured semantic repositories from relational databases // The semantic web: Research and applications. ESWC 2008. Lecture Notes in Computer Science*. – Springer, Berlin, Heidelberg, 2008. – T. 5021. – C. 777-781. doi: 10.1007/978-3-540-68234-9_57
25. Chungoora, N., Cutting-Decelle A.-F., Young R.I.M., Gunendran G., Usman Z., Harding J.A., Case K. *Towards the ontology-based consolidation of production-centric standards // International Journal of Production Research*. – 2013. – T. 51, № 2. – C. 327-345.
26. Compton M. et al. *Declarative Data Exchange: Spreadsheets to RDF/OWL // Proceedings of the 19th International Database Engineering & Applications Symposium*. – ACM, 2015. – C. 96-105.
27. 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*. – C. 491–494.
28. Cyganiak R. et al. *The D2RQ mapping language v0.8, 2012*. URL: <http://d2rq.org/d2rq-language>
29. 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). Lecture Notes in Computer Science*. – Tucson, AZ, USA, 2006. – T. 4215.– C. 241-254.

30. Erdmann M., Studer R. How to structure and access XML documents with ontologies // *Data & Knowledge Engineering*. – 2001. – T. 36, №. 3. – C. 317-335.
31. Gavrilova T., Leshcheva I., Blagov E., Yanson A., Bolotnikova E. Measuring Psychological Impact on Group Ontology Design and Development: an Empirical Approach // *Communications in Computer and Information Science*. – 2013. – T. 394. – C. 29-43.
32. Gómez-Pérez A. Some ideas and examples to evaluate ontologies // *Proceedings the 11th Conference on Artificial Intelligence for Applications*. – IEEE, 1995. – C. 299-305.
33. Gruber T. R. A translation approach to portable ontology specifications // *Knowledge acquisition*. – 1993. – T. 5. – №. 2. – C. 199-220.
34. Hacherouf M., Bahloul S. N., Cruz C. Transforming XML documents to OWL ontologies: A survey // *Journal of Information Science*. – 2015. – T. 41, №. 2. – C. 242-259.
35. Hevner A., Chatterjee S. *Design research in information systems: theory and practice*. – Springer Science & Business Media, 2010. – T. 22. – C. 5
36. Hu W., Qu Y. Discovering simple mappings between relational database schemas and ontologies // *The Semantic Web*. – Springer, Berlin, Heidelberg, 2007. – C. 225-238.
37. Madzík P., Budaj P., Mikuláš D., Zimon D. Application of the Kano Model for a Better Understanding of Customer Requirements in Higher Education — A Pilot Study // *Administrative Sciences*. – 2019. – T. 9, №. 1. – C. 11.
38. Nederstigt 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. – T. 59. – C. 296-311.
39. Österle H. et al. Memorandum on design-oriented information systems research // *European Journal of Information Systems*. – 2011. – T. 20, №. 1. – C. 7-10.
40. Otero-Cerdeira L., Rodríguez-Martínez F. J., Gómez-Rodríguez A. Ontology matching: A literature review // *Expert Systems with Applications*. – 2015. – T. 42, №. 2. – C. 949-971.
41. Ozturk O. OPPCAT: Ontology population from tabular data // *Journal of Information Science*. – 2019. – C.1-15. DOI: 10.1177/0165551519827892.
42. Paulheim H. Knowledge graph refinement: A survey of approaches and evaluation methods // *Semantic web*. – 2017. – T. 8, №. 3. – C. 489-508.
43. Peffers K. et al. A design science research methodology for information systems research // *Journal of management information systems*. – 2007. – T. 24, №. 3. – C. 45-77.
44. Petasis G. et al. Ontology population and enrichment: State of the art // *Knowledge-driven multimedia information extraction and ontology evolution*. – Springer-Verlag, 2011. – C. 134-166.
45. Rodrigues T., Rosa P., Cardoso J. Mapping XML to existing OWL ontologies // *International Conference WWW/Internet*. – 2006. – C. 72-77.
46. 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. – T. 11, №. 6. – C. 4456-4460.
47. 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. – C.27-34
48. Thuy P. T. T. et al. Exploiting XML schema for interpreting XML documents as RDF // *Services Computing, 2008. SCC'08. IEEE International Conference on*. – IEEE, 2008. – T. 2. – C. 555-558.

49. Tijerino Y. A., Embley D. W., Lonsdale D. W., Ding Y., Nagy G. Towards ontology generation from tables // World Wide Web. – 2005. – T. 8, №. 3. – C. 261-285.
50. Zhukova K., Pleshkova A., Mihnevich A., Pehtin I. One Approach to Administrative and Educational Processes Modeling: Case of Bachelor Program // International Journal of Education and Social Science. – 2014 - T. 3. – № 3. – C. 58-66.