

# СЕМАНТИЧЕСКИЙ ВЕБ И ГРАФЫ ЗНАНИЙ КАК ОБРАЗОВАТЕЛЬНАЯ ТЕХНОЛОГИЯ ПОДГОТОВКИ КАДРОВ ДЛЯ ЯДЕРНОЙ ЭНЕРГЕТИКИ

**В.П. Тельнов, Ю.А. Коровин**

*Обнинский институт атомной энергетики НИЯУ МИФИ  
249040, Россия, Калужская обл., г. Обнинск, Студгородок, 1*



Рассматриваются технологии представления знаний, модели рассуждений и методы генерации когнитивных гипотез в системах искусственного интеллекта. Практический акцент сделан на применении проблемно-ориентированных графов знаний как образовательной технологии при подготовке специалистов в области ядерной физики и атомной энергетики. Предлагаемые информационные решения и подходы охватывают задачи компьютерного обнаружения и классификации ядерных знаний и компетенций на основе онтологий, средства семантического аннотирования полнотекстового сетевого контента и представления информационных объектов в графовых базах данных, оснащенных средствами логического вывода. Прикладная часть проекта представлена общедоступным семантическим веб-порталом, который включает в себя прототипы следующих графов ядерных знаний: мировые центры ядерных данных, события и публикации ЦЕРН, базы данных и сетевые сервисы МАГАТЭ, учебные материалы МГУ и МИФИ по ядерной физике, ядерные исследовательские центры Российской Федерации, журналы по ядерной физике, объединенный граф ядерных знаний. Интерактивная визуальная навигация по графам знаний осуществляется с помощью специальных поисковых виджетов и интеллектуального браузера RDF. Браузер RDF позволяет пользователям совершать визуальные туры по графам знаний в любом направлении и на любое расстояние, по пути извлекая необходимые сведения в форме метаданных, гипертекстовых ссылок, полнотекстового и мультимедийного контента, ассоциированных с конкретным узлом графа. Операции с семантическими репозиториями выполняются на облачных платформах с использованием моделей обслуживания PaaS и DBaaS, что обеспечивает масштабируемость задействованных хранилищ данных и сетевых сервисов. Представлены результаты тестирования производительности семантической базы ядерных знаний и метрики вычислительных процессов. Рассматривается инновационный потенциал предлагаемых решений применительно к образовательной деятельности.

**Ключевые слова:** семантический веб, граф знаний, образовательный портал, онтология, графовая база данных, облачные вычисления.

## ВВЕДЕНИЕ И МОТИВАЦИЯ

Начиная с 60-х гг. XX в. в рамках исследований по искусственному интеллекту развивались формализмы для представления знаний [1] (семантические сети, фрей-

© В.П. Тельнов, Ю.А. Коровин, 2019

мовые системы и т.д.). В 2019 г. языки описания онтологий RDF [2], OWL [3], графы знаний и дескрипционные логики [4] дают современную теоретическую основу для создания систем и методов приобретения, представления, обработки и интеграции проблемно-ориентированных знаний в компьютерных системах, что, в частности, подтверждается действующими стандартами W3C [5] в области семантического веба. Актуальность семантических подходов отмечена в докладах III Международной конференции по управлению ядерными знаниями (2016 г.), в трудах международных конференций DAMDID/RCDL «Аналитика и управление данными в областях с интенсивным использованием данных» (2017, 2018 гг.). Семантические подходы к управлению знаниями находят применение при организации обмена научными данными (EUDAT); развитии международных баз знаний (DBpedia, Wikidata); при компьютерной обработке текстов, написанных на естественных языках (WordNet); семантическом анализе социальных сетей; в экспертных, прогностических системах, в медицине (IBM Watson).

В докладах International Conference on Semantic Systems [6], International Workshops on Description Logics [7] отмечается рост интереса гигантов IT-индустрии (Google, Facebook, Wikimedia) к графовым моделям представления знаний и к дескрипционным логикам. Попытки обнаружить в русскоязычном сегменте всемирной паутины действующие сервисы семантического веба редко приводят к успеху – в России производится недостаточно связанных открытых данных. Основными источниками данных для российских пользователей семантической паутины остаются общедоступные международные базы знаний, содержащие русскоязычный контент, в первую очередь DBpedia и Wikidata. По состоянию на 2019 г. образовательные веб-порталы университетов, центры ядерных данных, системы управления ядерными знаниями МАГАТЭ [8] и Росатома [9] не используют возможности семантической паутины в достаточной мере.

Целью рассматриваемого проекта является действующий семантический образовательный веб-портал нового поколения (WEB 3.0), который основан на стандартах и технологиях семантического веба и включает элементы искусственного интеллекта. Созданный авторами и представленный в статье прототип семантической базы ядерных знаний [10] поддерживает следующие графы:

- мировые центры ядерных данных;
- события и публикации ЦЕРН;
- базы данных и сетевые сервисы МАГАТЭ;
- учебные материалы МГУ и МИФИ по ядерной физике;
- ядерные исследовательские центры Российской Федерации;
- журналы по ядерной физике;
- объединенный граф ядерных знаний.

## **ПРЕДСТАВЛЕНИЕ ЗНАНИЙ, ОНТОЛОГИИ**

Онтология в контексте информационных технологий понимается как формальная спецификация с иерархической структурой, которая предназначена для представления знаний. Обычно онтология включает в себя описания классов сущностей (концептов, понятий) и свойств для некоторой предметной области, отношения между сущностями и ограничения на то, как эти отношения могут использоваться. Нередко онтологии воспринимаются как хранилища знаний, которые могут быть прочтаны и поняты как людьми, так и компьютерами, отчуждены от разработчика и повторно использованы.

На рисунке 1 в качестве примера представления знаний показан верхний уровень онтологии типа «Ядерный учебный центр», используемой в проекте [10] как основа для графа знаний «Учебные материалы МГУ и МИФИ по ядерной физике». В

данной онтологии структура классов сущностей близка к структуре учебно-методических комплексов дисциплин, преподаваемых в университетах. Онтологию, которая помимо классов сущностей и свойств включает в себя индивиды (экземпляры классов сущностей) и частные утверждения относительно индивидов, называют графом знаний. Можно сказать, что граф знаний – это онтология, обогащенная экстенциональными знаниями из конкретной предметной области. Богатство и разнообразие взаимоотношений между сущностями отличает онтологии и полноценные графы знаний от словарей, тезаурусов и таксономий, которые используются в консервативных системах управления знаниями (МАГАТЭ [8], Росатом [9], МИФИ [11]).

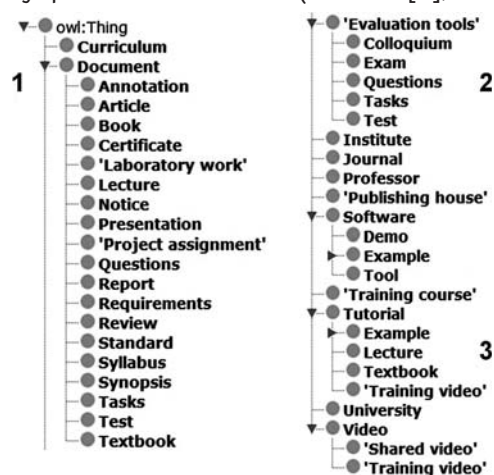


Рис. 1. Иерархия сущностей в онтологии типа «Ядерный учебный центр» (верхний уровень): 1 – виды документов; 2 – средства оценивания знаний; 3 – виды учебных пособий

На рисунке 2 в качестве примера показан фрагмент графа знаний «Учебные материалы МГУ и МИФИ по ядерной физике», содержащий семь классов и семь индивидов (экземпляров классов). Классы и индивиды связываются между собой при помощи различных свойств, которые имеют имена и направление, как видно из начертания и ориентации стрелок на рис. 2. Свойство «type» означает, что некий индивид принадлежит конкретному классу. Наименования большинства свойств очевидным образом раскрывают суть отношений между узлами в графе знаний, например, «is author of the lecture», «contains a textbook», «includes».

На рисунке 3 в качестве другого примера представлен фрагмент графа знаний «Мировые центры ядерных данных». Показан «Центр данных фотоядерных экспериментов» НИИ ядерной физики им. Д.В. Скобельцына, МГУ им. М.В. Ломоносова вместе с 12-ю базами ядерных данных и соответствующими онлайн-сервисами.

Возникает закономерный вопрос – какие новые возможности открывают графы знаний для специалистов, учащихся, преподавателей? Ответ дает интеллектуальный браузер RDF [2], который обеспечивает интерактивную визуальную навигацию по графам. Оказавшись в желаемом месте нужного графа знаний (например, при помощи поискового виджета), пользователь посредством этого браузера легко может выполнять визуальную навигацию по графу, посещая его узлы в нужном порядке и извлекая метаданные, гипертекстовые ссылки, полнотекстовый и медийный контент, ассоциированный с каждым узлом. При этом окрестность (окружение, замыкание) каждого узла графа становится видимой и доступной для навигации. Эта окрестность включает в себя узлы того графа, через который пользователь изначально вошел в семантическую паутину, а также смежные узлы иных графов, поддерживаемых базой знаний. Практически, процесс визуализации любых участков графа осуществляется простым щелчком компьютерной мыши по лепесткам, которые окружают узлы графа, или по пунктам локального меню.

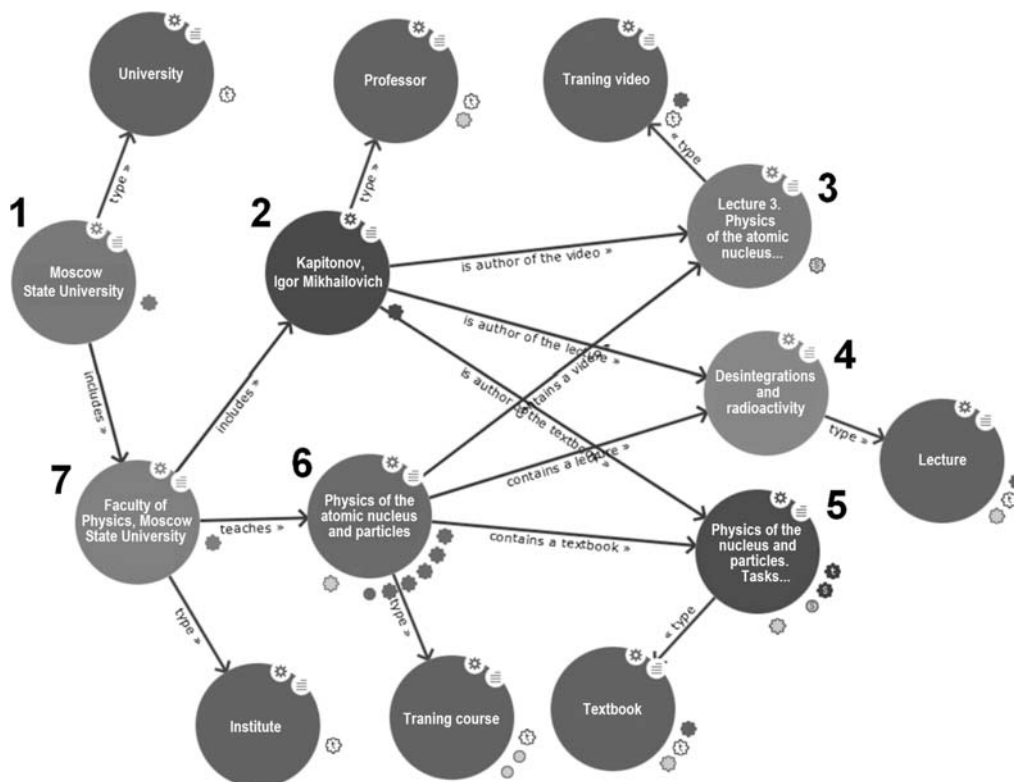


Рис. 2. Фрагмент графа знаний типа «Ядерный учебный центр» (пример): 1 – экземпляр класса «University»; 2 – экземпляр класса «Professor»; 3 – экземпляр класса «Training video»; 4 – экземпляр класса «Lecture»; 5 – экземпляр класса «Textbook»; 6 – экземпляр класса «Training course»; 7 – экземпляр класса «Institute»

На рисунке 3 показаны ближайшая окрестность узла, который соответствует индивиду «Photonuclear Experiment Data Center» класса «Laboratory», а также некоторые элементы управления, ассоциированные с узлами графа. Если сфокусироваться на любом ином узле, который отображается браузером RDF, он также станет доступен для навигации вместе с его окрестностью и метаданными. Таким образом, пользователь имеет возможность осуществлять произвольные визуальные туры по графам знаний в любом направлении и на любую глубину, выбирая появляющиеся данные. При наведении курсора на различные участки узлов графа появляются локальное меню, уточняющая информация и подсказки для вариантов продолжения навигации по графу.

Графы знаний могут физически размещаться в удаленных графовых базах данных или в иных репозиториях, именуемых RDF-хранилищами или хранилищами триплетов [12]. Каждый из графов знаний содержит тысячи триплетов. Для быстрого погружения в доступные графы знаний в проекте [10] пользователям предлагаются поисковые виджеты, которые позволяют попадать в нужное место конкретного графа, где будут обнаружены и визуализированы искомые информационные объекты. Принцип работы поисковых виджетов аналогичен тому, как происходит выборка информации из всемирной паутины при помощи популярных поисковых систем (Google, Яндекс и т.д.). По мере того как пользователь набирает символы ключевых слов в строке ввода поискового виджета, система «выкачивает» адекватный список сущностей из соответствующего графа знаний. Пользователю предлагается выбрать подходящий класс или индивид и погрузиться непосредственно в нужную область графа. Далее становится возможной более точная визуальная навигация по

графу, которая осуществляется интуитивно понятным образом с применением интеллектуального браузера RDF.

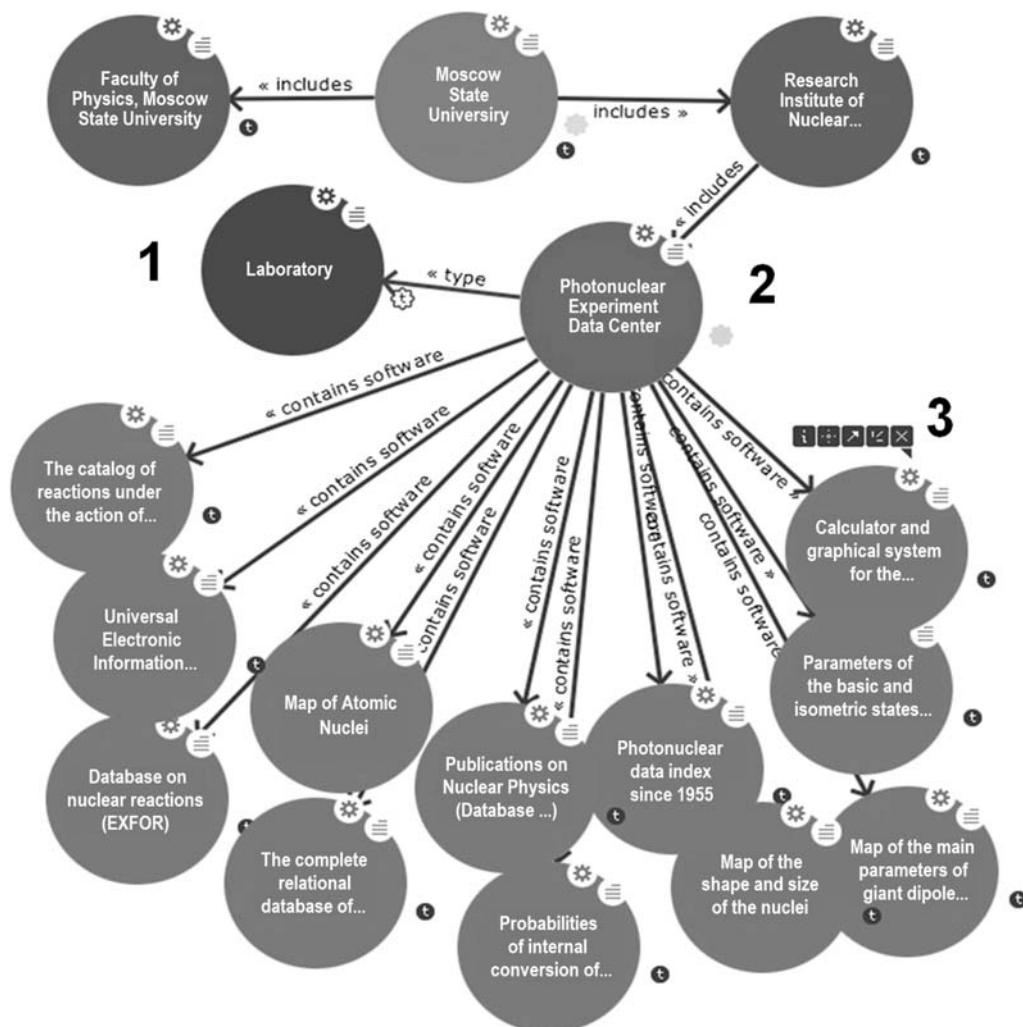


Рис. 3. Фрагмент графа знаний типа «Мировые центры ядерных данных» (пример): 1 – класс «Laboratory»; 2 – «Photonuclear Experiment Data Center», экземпляр класса «Laboratory»; 3 – всплывающее локальное меню для навигации по графу знаний

### ЛОГИЧЕСКИЕ РАССУЖДЕНИЯ НА ГРАФАХ ЗНАНИЙ

Одна из привлекательных особенностей семантического веба заключается в том, что становится возможным извлечение (вывод) новых знаний из фактов, которые уже имеются в графе знаний. Для этого применяются интеллектуальные программные агенты, называемые ризонерами [13]. То, как логический вывод осуществляется алгоритмически, не специфицируется в самой онтологии (графе знаний) или в соответствующем OWL-документе, поскольку OWL – это декларативный язык описания онтологий. Правильный ответ на любой вопрос предопределяется семантикой той дескрипционной логики, которая задает стандарт языка. Особый интерес представляет способность баз знаний как систем искусственного интеллекта генерировать когнитивные гипотезы, используя для этого помимо дедукции и иные способы рассуждений, такие как индуктивные рас-



суждения [14], аргументация, рассуждения на основе прецедентов.

Что касается индуктивных правил вывода в графах, их целесообразность обусловлена следующими соображениями. Во-первых, индуктивные правила вывода на основе рассмотрения возможных альтернатив (прецедентов) позволяют генерировать когнитивные гипотезы (нечеткие знания), которые не могут быть получены непосредственно при дедуктивных рассуждениях на графе. Во-вторых, индуктивный вывод [15] – это одна из базовых технологий семантического аннотирования сетевого контента, когда необходимо конструировать, расширять и актуализировать имеющиеся графы новыми знаниями. Отыскание и селекция адекватных сетевых информационных ресурсов в проекте [10] осуществляется посредством контекстно-зависимого разведочного поиска, соответствующая технология предложена в [16]. При помощи индуктивного вывода решаются задачи классификации и кластеризации новых информационных объектов в семантической базе ядерных знаний.

Дескрипционные логики [4] позволяют описывать различные области знаний в недвусмысленном формализованном виде. Они сочетают в себе, с одной стороны, богатые выразительные возможности, а с другой – хорошие вычислительные свойства, такие как разрешимость и относительно невысокая вычислительная сложность при решении основных задач логического вывода. Дескрипционные логики можно рассматривать как разрешимые фрагменты логики исчисления предикатов первого порядка, когда не приемлема экспоненциальная, но допустима полиномиальная сложность вычислений. С точки зрения дескрипционных логик, набор утверждений общего вида, или терминология, обозначается как TBox (интенциональное знание). TBox формирует онтологию в собственном смысле этого слова. Также выделяют наборы утверждений частного вида, которые обозначаются как ABox (экстенциональное знание). TBox совместно с ABox образуют онтологию (граф знаний) в широком понимании этого слова. Графы знаний проекта [10] обслуживаются визуальным ризонером, который встроен в интеллектуальный браузер RDF и обеспечивает следующие виды стандартных проверок и рассуждений на графах:

- при проверке выполнимости (satisfiability) исследуется вопрос, может ли произвольный класс иметь экземпляры (индивиды);
- при проверке категоризации (subsumption) исследуются подклассы классов, строится иерархия вложенности классов и свойств;
- при проверке согласованности (consistency) исследуются вопрос, насколько экстенциональные знания (ABox) согласованы в графе с интенциональными знаниями (TBox);
- при проверке экземпляров классов (индивидов) исследуется корректность всех утверждений о данном экземпляре с точки зрения ABox;
- поиск всех экземпляров класса (индивидов), а также поиск всех классов, которым конкретный индивид принадлежит прямо или опосредованно;
- поиск всех сущностей, которые связаны между собой определенным свойством, а также группировка (кластеризация) таких сущностей.

Примеры автоматизированных логических рассуждений на графах ядерных знаний проекта [10] приведены в [17]. Кластеры сущностей, связанные между собой определенным свойством или группой свойств, – это примеры выведенных фактов (образцы новых знаний), которые изначально в явном виде не были представлены в графе. Выведенные факты в браузере RDF имеют вид лепестков, сгруппированных вокруг узлов графа; они раскрываются щелчком мыши и весьма удобны для последующей визуальной навигации по графу. Вообще, визуальный способ указания правил вывода на графе знаний является своего рода «кизюминкой» проекта [10], которая отличает его от традиционных ризонеров, где правила вывода задаются с использованием логических

предикатов и формального синтаксиса [18]. Представляется, что интуитивно понятный визуальный способ задания правил вывода и навигации является комфортным для малоискушенных пользователей графов знаний.

### ОЦЕНКА ПРОИЗВОДИТЕЛЬНОСТИ ГРАФОВ ЗНАНИЙ

Обсуждаемый проект [10] открыт для свободного использования неограниченным кругом заинтересованных лиц, однако следует отметить два обстоятельства: 1) на дату подготовки статьи работает пилотная версия семантической базы ядерных знаний, наполнение графов знаний продолжается; 2) серверная часть программного обеспечения функционирует на облачных платформах по бесплатной квоте, поэтому возможны задержки с обслуживанием запросов.

Вычислительные эксперименты проводились на графах знаний «Мировые центры ядерных данных», «Учебные материалы МГУ и МИФИ по ядерной физике» и международной базе знаний «DBpedia» по разным сценариям (табл. 1).

Таблица 1  
**Результаты тестирования производительности семантической базы ядерных знаний**

Вид теста	Тестовые данные	Метрики вычислительного процесса
Запуск главной страницы семантического веб-портала (поисковые виджеты)	Граф знаний «Мировые центры ядерных знаний»	Количество сетевых запросов: 162. Загрузка веб-страницы: 1140 мс. Загрузка репозитория: 850 мс. Отработка серверного скрипта: 150 мс.
Загрузка и визуализация графа знаний (браузер RDF)	Граф знаний «Мировые центры ядерных знаний», объект «Photonuclear Experiment Data Center»	Количество сетевых запросов: 39. Загрузка веб-страницы: 893 мс. Загрузка репозитория: 402 мс. Отработка серверного скрипта: 44 мс.
Загрузка и визуализация графа знаний (браузер RDF)	Граф знаний «Учебные материалы МГУ и МИФИ по ядерной физике», объект «Physics of the atomic nucleus and particles»	Количество сетевых запросов: 57. Загрузка веб-страницы: 1050 мс. Загрузка репозитория: 643 мс. Отработка серверного скрипта: 17 мс.
Загрузка и визуализация международной базы знаний «DBpedia», (браузер RDF)	Международная база знаний «DBpedia», объект «World War II»	Количество сетевых запросов: 37. Загрузка веб-страницы: 2060 мс. Загрузка репозитория: 753 мс. Отработка серверного скрипта: 15 мс.

Метрики вычислительных процессов, представленные в таблице, были получены при следующих условиях испытаний:

- семантические репозитории физически размещаются на облачной платформе Amazon Web Service [19] по бесплатной квоте, центр обработки данных расположен в Западной Европе;
- каждый граф знаний обслуживается выделенным виртуальным компьютером;
- измеренная скорость интернет-соединения составляет около 90 Мбит/с;
- в качестве клиентского компьютера используется рабочая станция с процессором Intel Core i5-8400 2,8/4,0 ГГц и памятью 16 МБ, расположенная в г. Обнинске;
- эксперименты проводились на тестовых онтологиях, которые включали в себя не более тысячи сущностей (DBpedia – исключение).

Расходование вычислительного ресурса детализируется в отладочных панелях браузеров (Google Chrome, Mozilla Firefox), соответствующие параметры можно наблюдать вживую при работе с семантическим веб-порталом [10]. Во всех экспериментах общее время отработки не превышало двух – трех секунд, при этом большая часть времени и подавляющая доля вычислительного ресурса расходовались на

функционирование интерфейса к базе знаний и на сетевой трафик. Пример онтологии «Учебные материалы МГУ и МИФИ по ядерной физике» в сериализованном формате доступен по ссылке [20]. Архитектура программного обеспечения проекта представлена в [17].

### **ЗАКЛЮЧЕНИЕ**

Потенциальные бенефициары информационных решений и технологий, которые обозначены как результаты проекта [10] (целевая аудитория), – студенты, преподаватели, руководители, эксперты, исследователи и специалисты в области ядерной физики, атомной энергетики, ядерной и радиационной безопасности. Предлагаемые программные решения базируются на облачных вычислениях по моделям обслуживания PaaS и DBaaS, что обеспечивает масштабируемость хранилищ данных и сетевых служб.

Рассмотрим инновационный потенциал проекта [10] применительно к образовательной деятельности. Студенты и преподаватели университетов расходуют немало времени и усилий на поиск полезной информации во всемирной паутине вместо того, чтобы воспринимать и интерпретировать достоверный учебный материал. Современные реалии высшего образования таковы, что значительное число студентов и преподавателей не подозревают о существовании семантического веба и связанных открытых данных. В учебном процессе они используют традиционные системы управления контентом (CMS), системы управления обучением (LMS) и виртуальные учебные среды (VLE), которые основаны на таксономиях и тезаурусах (Coursera, edX, Moodle и т.д.). Студенты широко практикуют поиск информации во всемирной паутине по ключевым словам, используя для этого общедоступные поисковые системы, при этом они усваивают значительное количество недостоверных сведений, рекламы и информационного мусора. Существенную роль здесь играют сформировавшиеся традиции, а также легкость и большая скорость формирования поисковых запросов к публичным поисковым системам по сравнению с запросами к семантической паутине. При этих обстоятельствах выглядит оправданным известный скептицизм относительно того, что семантические интернет-технологии в их нынешнем виде получают признание и широкое распространение в университетской и в профессиональной среде. Можно предположить, что коммерциализированные поисковые системы, наряду с различными вариантами википедий, еще долго будут оставаться наиболее востребованными «универсальными учебниками» для многочисленных студентов и специалистов, если на смену им не станут приходиться полноценные графы знаний и интеллектуальные поисковые агенты, снабженные адекватными средствами визуализации.

### **Благодарности**

Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов № 18-07-00583 а, № 19-47-400002 р\_а, фонда В. Потанина проект № ГК190001383.

### **Литература**

1. *Harmelen F., Lifschitz V., Porter B.* Handbook of Knowledge Representation. England, Oxford: Elsevier Science Oxford, 2008, ISBN 78-0-444-52211-5.
2. W3C RDF Schema 1.1. Электронный ресурс: <http://www.w3.org/TR/rdf-schema> (дата доступа 25.03.2019).
3. W3C OWL 2 Web Ontology Language. Электронный ресурс: <http://www.w3.org/TR/owl2-overview> (дата доступа 25.03.2019).
4. *Baader F., Calvanese D., McGuinness D., Nardi D., Patel-Schneider P.* The Description Logic



Handbook: Theory, Implementation and Applications, 2-nd Ed. New York, USA, Cambridge University Press, 2010, ISBN 0521150116 9780521150118.

5. W3C Semantic Web. Электронный ресурс: <http://www.w3.org/standards/semanticweb> (дата доступа 25.03.2019).

6. Proceedings of the 14th International Conference on Semantic Systems. Vienna, Austria, 10-13 September 2018. Электронный ресурс: <http://www.sciencedirect.com/journal/procedia-computer-science/vol/137> (дата доступа 25.03.2019).

7. Proceedings of the 31st International Workshop on Description Logics. Tempe, Arizona, US, 27-29 October, 2018. Электронный ресурс: <http://ceur-ws.org/Vol-2211/> (дата доступа 25.03.2019).

8. Knowledge Management for Nuclear Research and Development Organizations, IAEA TECDOC No. 1675. Электронный ресурс: <http://www-pub.iaea.org/books/IAEABooks/8644/Knowledge-Management-for-Nuclear-Research-and-Development-Organizations> (дата доступа 25.03.2019).

9. Лещенко В. Управление знаниями Росатома. История успеха. Электронный ресурс: [http://www.imodel-russia.com/upload/MI\\_dokladi/MI\\_Rosatom\\_Leshchenko\\_NEOLANT\\_Kutilin.pdf/](http://www.imodel-russia.com/upload/MI_dokladi/MI_Rosatom_Leshchenko_NEOLANT_Kutilin.pdf/) (дата доступа 25.03.2019).

10. Семантическая база ядерных знаний. Электронный ресурс: <http://vt.obninsk.ru/x/> (дата доступа 25.03.2019).

11. Образовательный портал НИЯУ МИФИ. Электронный ресурс: <http://online.mephi.ru/> (дата доступа 25.03.2019).

12. Howard P. Graph and RDF databases. Market Report Paper by Bloor 2018. Электронный ресурс: <http://www.bloorresearch.com/technology/graph-databases> (дата доступа 25.03.2019).

13. The University of Manchester. List of Reasoners. Электронный ресурс: <http://owl.cs.manchester.ac.uk/tools/list-of-reasoners> (дата доступа 25.03.2019).

14. D'Amato C., Fanizzi N., Fazzino B., Gottlob G., Lukasiewicz T. Combining Semantic Web Search with the Power of Inductive Reasoning (2013). Электронный ресурс: <http://ceur-ws.org/Vol-527/paper2.pdf/> (дата доступа 25.03.2019).

15. Minervini P., d'Amato C., Fanizzi N., Tresp V., Discovering Similarity and Dissimilarity Relations for Knowledge Propagation in Web Ontologies. // Journal on Data Semantics. – 2016. – Vol. 5. – Iss. 4, December. – PP. 229-248.

16. Тельнов В.П. Контекстный поиск как технология извлечения знаний в сети Интернет. // Программная инженерия. – 2017. – № 1. – С. 26-37.

17. Тельнов В.П., Коровин Ю.А. Программирование графов знаний, рассуждения на графах // Программная инженерия. – 2019. – № 2. – С. 59-68.

18. Reasoners and rule engines: Jena inference support. – Электронный ресурс: <http://jena.apache.org/documentation/inference> (дата доступа 25.03.2019).

19. Amazon Web Services. Электронный ресурс: <http://aws.amazon.com/ru> (дата доступа 25.03.2019).

20. Пример онтологии «Учебные материалы МГУ и МИФИ по ядерной физике». Электронный ресурс: <http://drive.google.com/file/d/1AIXMsm3cfAxR6NX220R4ZeFeoSfp0mj5> (дата доступа 25.03.2019).

Поступила в редакцию 11.04.2019 г.

#### Авторы

Тельнов Виктор Петрович, доцент, канд. техн. наук  
E-mail: [telnov@bk.ru](mailto:telnov@bk.ru)

Коровин Юрий Александрович, профессор, д-р физ.-мат. наук  
E-mail: [korovinyu@mail.ru](mailto:korovinyu@mail.ru)

## SEMANTIC WEB AND KNOWLEDGE GRAPHS AS AN EDUCATIONAL TECHNOLOGY OF PERSONNEL TRAINING FOR NUCLEAR POWER ENGINEERING

Telnov V.P., Korovin Yu.A.

Obninsk Institute for Nuclear Power Engineering NRNU MEPhI  
1 Studgorodok, Obninsk, Kaluga reg., 249040 Russia

### ABSTRACT

The paper is devoted to knowledge representation technologies, reasoning models and methods of generating cognitive hypotheses in artificial intelligence systems. Practical emphasis is placed on the use of problem-oriented knowledge graphs as an educational technology in the training of specialists in the field of nuclear physics and nuclear energy. The proposed information solutions and approaches cover the tasks of computer detection and classification of nuclear knowledge and competences on the basis of ontologies, semantic annotation tools for full-text network content and representation of information objects in graph databases equipped with logical inference tools. The applied part of the project are presented publicly available semantic web portal which includes prototypes for the following graphs of nuclear knowledge: the world centre of nuclear data; events and publications CERN; databases and network services of the IAEA; educational materials of the MSU and MEPhI in nuclear physics; nuclear research centers of the Russian Federation; journals for nuclear physics; the joint graph of nuclear knowledge. Interactive visual navigation through knowledge graphs is performed using special search widgets and an intelligent RDF browser. The RDF browser allows users to make visual tours of knowledge graphs in any direction and at any distance, extracting the necessary information in the form of metadata, hypertext links, full-text and media content associated with a particular node of the graph. Operations with semantic repositories are performed on cloud platforms using PaaS and DBaaS service models, which provides scalability of the involved data stores and network services. The results of performance testing of the semantic base of nuclear knowledge and metric of the computational processes are presented. The innovative potential of the proposed solutions in relation to educational activities is considered.

**Key words:** semantic web, educational portal, ontology, graph database, knowledge graph, cloud computing.

### REFERENCES

1. Harmelen F., Lifschitz V., Porter B. *Handbook of Knowledge Representation*. England, Oxford: Elsevier Science Oxford, 2008, ISBN 78-0-444-52211-5.
2. W3C RDF Schema 1.1. Available at: <http://www.w3.org/TR/rdf-schema> (accessed Mar 25, 2019).
3. W3C OWL 2 Web Ontology Language. Available at: <http://www.w3.org/TR/owl2-overview> (accessed Mar 25, 2019).
4. Baader F., Calvanese D., McGuinness D., Nardi D., Patel-Schneider P. *The Description Logic Handbook: Theory, Implementation and Applications*, 2nd Ed. New York, USA, Cambridge University Press, 2010, ISBN 0521150116 9780521150118.
5. W3C Semantic Web. Available at: <http://www.w3.org/standards/semanticweb> (accessed Mar 25, 2019).
6. Proceedings of the 14th International Conference on Semantic Systems. Vienna, Austria, 10–13 September 2018. Available at: <http://www.sciencedirect.com/journal/>

procedia-computer-science/vol/137 (accessed Mar 25, 2019).

7. Proceedings of the 31st International Workshop on Description Logics. Tempe, Arizona, US, 27-29 October, 2018. Available at: <http://ceur-ws.org/Vol-2211/> (accessed Mar 25, 2019).

8. Knowledge Management for Nuclear Research and Development Organizations, IAEA TECDOC No. 1675. Available at: <http://www-pub.iaea.org/books/IAEABooks/8644/Knowledge-Management-for-Nuclear-Research-and-Development-Organizations> (accessed Mar 25, 2019).

9. Leshchenko V. Rosatom knowledge management. History of success. Available at: [http://www.imodel-russia.com/upload/MI\\_dokladi/MI\\_Rosatom\\_Leshchenko\\_NEOLANT\\_Kutilin.pdf/](http://www.imodel-russia.com/upload/MI_dokladi/MI_Rosatom_Leshchenko_NEOLANT_Kutilin.pdf/) (accessed Mar 25, 2019) (in Russian).

10. The semantic Base of Nuclear Knowledge. Available at: <http://vt.obninsk.ru/x/> (accessed Mar 25, 2019) (in Russian).

11. Educational portal of NRNU MEPhI. Available at: <http://online.mephi.ru/> (accessed Mar 25, 2019) (in Russian).

12. Howard P. Graph and RDF databases. Market Report Paper by Bloor 2018. Available at: <http://www.bloorresearch.com/technology/graph-databases> (accessed Mar 25, 2019).

13. The University of Manchester. List of Reasoners. Available at: <http://owl.cs.manchester.ac.uk/tools/list-of-reasoners> (accessed Mar 25, 2019).

14. D'Amato C., Fanizzi N., Fazzinga B., Gottlob G., Lukasiewicz T. Combining Semantic Web Search with the Power of Inductive Reasoning (2013). Available at: <http://ceur-ws.org/Vol-527/paper2.pdf> (accessed Mar 25, 2019).

15. Minervini P., d'Amato C., Fanizzi N., Tresp V., Discovering Similarity and Dissimilarity Relations for Knowledge Propagation in Web Ontologies. *Journal on Data Semantics*. 2016, v. 5, iss. 4, December, pp. 229-248.

16. Telnov V.P. The Contextual Search as a Technique for Extracting Knowledge on the Internet. *Programmnyaya Inzheneriya*. 2017, no. 1, pp. 26-37 (in Russian).

17. Telnov V.P., Korovin Yu.A. Programming of Knowledge Graphs, Reasoning on Graphs. *Programmnyaya Inzheneriya*. 2019, no. 2, pp. 59-68 (in Russian).

18. Reasoners and rule engines: Jena inference support. Available at: <http://jena.apache.org/documentation/inference> (accessed Mar 25, 2019).

19. Amazon Web Services. Available at: <http://aws.amazon.com/ru> (accessed Mar 25, 2019).

20. An example of ontology «Educational materials of Moscow State University and NRNU MEPhI on Nuclear Physics». Available at: <http://drive.google.com/file/d/1AIXMsm3cfAxR6NX220R4ZeFeoSfp0mj5> (accessed Mar 25, 2019) (in Russian).

#### Authors

Telnov Victor Petrovich, Associate Professor, Cand. Sci. (Engineering)

E-mail: [telnov@bk.ru](mailto:telnov@bk.ru)

Korovin Yuri Alexandrovich, Professor, Dr. Sci. (Phys.-Math.)

E-mail: [korovinyu@mail.ru](mailto:korovinyu@mail.ru)