



УДК 004.043

ИССЛЕДОВАНИЕ АКТУАЛЬНЫХ СПОСОБОВ ПУБЛИКАЦИИ ОТКРЫТЫХ НАУЧНЫХ ДАННЫХ В СЕТИ

Д.И. Муромцев^a, Й. Леманн^b, И.А. Семерханов^a, М.А. Навроцкий^a, И.С. Ермилов^b^a Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация^b Лейпцигский Университет, Лейпциг, 04109, Германия

Адрес для переписки: m.navrotskiy@gmail.com

Информация о статье

Поступила в редакцию 19.09.15, принята к печати 17.10.15

doi:10.17586/2226-1494-2015-15-6-1081-1087

Язык статьи – русский

Ссылка для цитирования: Муромцев Д.И., Леманн Й., Семерханов И.А., Навроцкий М.А., Ермилов И.С. Исследование актуальных способов публикации открытых научных данных в сети // Научно-технический вестник информационных технологий, механики и оптики. 2015. Том 15. № 6. С. 1081–1087.

Аннотация

Предмет исследования. Приведен обзор программных средств и технологий публикации открытых данных в сети Интернет в машиночитаемых форматах в контексте данных университетов, образовательных и исследовательских организаций, научных лабораторий. Выполнен анализ наиболее часто используемых форматов публикации, включая не только такие популярные форматы, как pdf, csv, excel, но и форматы Семантического Веба, например, RDF. Рассмотрены способы публикации научных данных в семантических форматах на примере импорта и конвертации информации из базы данных университета. **Методы.** Описаны методы публикации открытых научных данных в сети, состоящие из набора преобразований исходных массивов данных к конечному семантическому представлению. Показаны этапы преобразований, включая выгрузку данных из реляционной базы данных, отображение (маппинг) данных на онтологическую модель (схему) и генерацию массива RDF-триплетов, соответствующего исходному фрагменту базы данных. Приведено описание популярных систем публикации открытых данных, таких как SKAN, VIVO и другие. В качестве основной системы хранения и публикации данных рассмотрена система OpenLinkVirtuoso. Дано описание модели данных RDF в качестве способа представления открытых данных университета ИТМО. **Основные результаты.** Приведено описание способов публикации открытых научных данных, показаны их недостатки. Для демонстрации работоспособности предложенного способа публикации открытых данных университета разработан программный прототип, доступный в сети Интернет. Приведен пример использования системы. **Практическая значимость.** Внедрение предложенного подхода позволит значительно повысить эффект от публикации открытых данных университетов для использования этих данных другими приложениями, например, приложениями поиска информации об образовательной деятельности и результатах исследований, анализа научной деятельности университетов и их научных подразделений.

Ключевые слова

онтологии, RDF, связанные открытые данные, интеграция данных, публикация данных, virtuoso, sparql.

STUDY OF CURRENT APPROACHES FOR WEB PUBLISHING OF OPEN SCIENTIFIC DATA

D.I. Mouromtsev^a, J. Lehmann^b, I.A. Semerkhanov^a, M.A. Navrotskiy^a, I.S. Ermilov^b^a ITMO University, Saint Petersburg, 197101, Russian Federation^b University of Leipzig, Leipzig, 04109, Germany

Corresponding author: m.navrotskiy@gmail.com

Article info

Received 19.09.15, accepted 17.10.15

doi:10.17586/2226-1494-2015-15-6-1081-1087

Article in Russian

For citation: Mouromtsev D.I., Lehmann J., Semerkhanov I.A., Navrotskiy M.A., Ermilov I.S. Study of current approaches for Web publishing of open scientific data. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2015, vol. 15, no. 6, pp. 1081–1087.

Abstract

Subject of Study. The subject of study of this work is closely related to the development of tools and technologies for Internet publishing of open data in machine-readable formats with regard to data of universities, educational and research organizations and scientific laboratories. We analyze the trends in the publishing formats most commonly used including not only popular formats such as pdf, csv, excel, but also the Semantic Web formats such as RDF. The paper describes the way of scientific data publication in semantic formats on the example of import and conversion of the information from University database. **Methods.** We describe the methods of publication for scientific open data in the network consisting of a

set of transformations of the original data sets to the final semantic representation. These transformation steps include data upload from a relational database, data mapping on the ontological model (schema) and the generation of a set of RDF-triples corresponding to the initial database fragment. A description is given to the popular open data publishing systems, such as CKAN, VIVO, and others. OpenLink Virtuoso system is selected as the primary storage and data publication. The description of RDF data model is used as a way of presenting open data of ITMO University. **Main Results.** The authors have described the methods of scientific open data publication and identified their shortcomings. To demonstrate the efficiency of the proposed method of university open data publication, a software prototype has been developed available online at: <http://lod.ifmo.ru/>. The example of the system usage is also given. **Practical Relevance.** Implementation of the proposed approach will improve significantly the effect of the publication of university open data and make it available for third-party applications, such as applications for information retrieval about educational activities and research results, analysis of scientific activities in universities and their research departments.

Keywords

ontology, RDF, linked open data, data integration, data publishing, virtuoso, sparql.

Введение

Все больший объем данных публикуется в открытом виде в сети. Правительства, государственные администрации, научные центры, лаборатории, университеты, коммерческие компании реализуют открытые хранилища в широком диапазоне секторов: расходы, преступность, образование, здравоохранение, транспорт, окружающая среда. Примерами являются порталы открытых данных США¹, Великобритании², Канады³, ЕС⁴, и многие другие. Эти порталы публикуют данные во многих форматах, но в основном используются CSV или Excel, очень небольшое количество данных публикуется в виде RDF⁵, несмотря на его преимущества. Это можно объяснить тем фактом, что публикация данных в формате RDF требует специальных знаний и больших трудозатрат по сравнению с традиционными решениями. Множество таких примеров предоставляет Веб научных данных [1].

Несмотря на то, что объем открытой информации в Интернете растет на 57% в год [2], Веб научных данных сохраняет значительную инерцию, его рост колеблется между 1,6% и 14% [2] в зависимости от типа публикации и области исследования. Доля Веба научных данных во всем Вебе крайне мала.

Ситуацию можно изменить за счет публикации университетами результатов о своей научной деятельности. Однако данные университетов гораздо чаще публикуются в привычных и простых форматах, как, например, PDF, а не в формате RDF. Для преобразования уже опубликованных результатов в нужный формат необходимо использовать специальные утилиты.

В статье рассматривается проблема извлечения данных из реляционных баз Университета ИТМО. В данной области разработано целое семейство языков, таких как W3C R2RML [3], D2RQ [4], R2O [5], и их реализации, например, в VirtuosoUniversal Server [6]. Однако по-прежнему является большой проблемой механизм автоматического установления соответствия между реляционной базой данных (БД) и RDF-структурой.

Язык R2RML используется в основном для преобразования SQL-схем в RDF, если конечный RDF-граф имеет схожую с БД структуру и не сильно отличается от нее. Он представляет собой набор SQL-запросов на выборку из БД, которые также содержат информацию о том, каким образом результаты должны быть преобразованы в формат RDF [3].

Язык D2RQ использует отличный от R2RML подход. Вместо создания RDF-дамп-файла создается файл соответствия RDF-структуры реляционной БД. В момент выполнения SPARQL-запроса он трансформируется в SQL-запрос. Другое преимущество D2RQ заключается в том, что существует возможность задавать отображение (маппинг БД) по умолчанию с последующей возможностью кастомизации, что упрощает начальные шаги по трансформации. К недостаткам можно отнести высокое время выполнения SPARQL-запроса, а в случае меняющихся БД соответствие трудно поддерживать [4].

Язык R2O похож на D2RQ, но, в отличие от предыдущего, преобразует данные в OWL-формат [5].

Таким образом, для решения поставленной задачи необходимо спроектировать и разработать специализированные приложения, которые, с одной стороны, облегчат публикацию научных данных в RDF, а с другой – позволят преобразовать уже опубликованные данные в нужный формат из других форматов.

В настоящей работе дано описание прототипа программной системы, приведено описание используемых технологий и результаты исследований существующих способов публикации открытых связанных данных. Показан пример использования предложенного решения. Для извлечения данных из реляционных БД использованы язык R2RML и программное средство sparqlify⁶.

¹ <http://www.data.gov>

² <http://data.gov.uk>

³ <http://open.canada.ca/en>

⁴ <https://open-data.europa.eu/en/data/>

⁵ <http://www.w3.org/RDF/>

⁶ <http://sparqlify.org/>

Связанные открытые научные данные

Семантический Веб (Semantic Web) – это Веб машиночитываемых данных, где каждый источник данных может использоваться в различных приложениях и системах [7]. Одним из лучших решений для публикации взаимосвязанных данных [8] является использование технологии связанных данных (Linked Data), основанной на стандартах W3C и рекомендуемых технологиях, таких как RDF и OWL.

В области публикации научной информации в открытых форматах достаточно распространены, например, такие платформы, как GoogleScholar¹, Academia.eu², BingAcademic [9], но опубликованные в них материалы не являются семантическими и несовместимы с RDF. Данные в таких платформах публикуются в формате PDF и нуждаются в дополнительной аннотации для связи с ранее опубликованными материалами, что увеличивает время на их поиск и анализ.

Университеты, бесспорно, играют ключевую роль в развитии этого направления. В настоящий момент в ведущих университетах Европы и США уже существует несколько проектов, использующих принципы Semantic Web и связанных данных:

- Университет Мюнстера поддерживает проект LODUM – портал, содержащий различную информацию о научной и образовательной деятельности в машиночитаемом формате RDF;
- Оксфордский университет достаточно давно использует для публикации результатов своей деятельности открытое хранилище <http://data.ox.ac.uk>;
- Университет Саутгемптона предоставляет доступ к своему portalу открытых данных <http://data.southampton.ac.uk>;
- Университет Бонна разрабатывает открытую платформу для публикации открытой информации <http://opendata.bonn.de>;
- Гарвардский университет предоставляет доступ к публикациям в открытом виде при помощи инструмента HarvardDataverse <https://dataverse.harvard.edu>.

Использование связанных данных в научной среде открывает новые возможности по поиску информации одновременно во многих источниках, таких как публикации, научные сотрудники, научные результаты, лаборатории.

Способы публикации открытых научных данных

В описанных проектах для публикации научной информации преимущественно использованы системы VIVO [10, 11], SKAN и некоторые другие. Рассмотрим их более подробно.

- VIVO – это стек технологий для построения междисциплинарной сети. Проект выделяется среди других систем из-за его широкого внедрения в научное сообщество и использования семантических Веб-технологий. Первоначально он был разработан для интеграции данных внутри Университета Корнелл в 2006 году. Проект VIVO был продлен в 2009 году для поддержки кросс-университетской интеграции данных [11].
- SKAN предоставляет метаданные о данных в каталоге, позволяет публиковать, искать и использовать их. SKAN предоставляет пользователям и разработчикам средства для легкой публикации данных. Опубликованная информация может быть изучена конечным потребителем посредством свободного поиска и поиска на основе различных атрибутов, групп наборов и тегов.
- ResearchGate – это приложение для международного сообщества исследователей. Оно поддерживает более широкое распространение публикаций в формате PDF, предоставляя к ним доступ более чем 2 000 000 пользователям.
- Dataverse – это проект Университета Гарварда, представляющий собой Веб-приложение для публикации и анализа исследовательских результатов, которое базируется на пользовательских решениях для хранения данных.
- Ambra – система публикации научных статей и журналов, которая поддерживает пост-публикационную аннотацию для соотнесения различных публикаций.

Хотя системы, публикующие открытые данные из различных источников, уже существуют, отсутствуют единая платформа и методология, которые могут применяться в любом университете или исследовательском центре для публикации результатов исследований.

Проблемы существующих способов

Все вышеописанные системы успешно используются, однако их ключевая проблема заключается в невозможности автоматической обработки опубликованных данных. Одним из перспективных решений этой проблемы является внедрение процесса экстернализации при публикации данных. Экстернализация – процесс формулирования неявных знаний посредством явных понятий [12]. Идея этого процесса

¹ <http://scholar.google.ru>

² <http://www.academia.edu>

состоит в явном представлении данных научных результатов и обеспечении доступа к ним в сети Интернет таким образом, чтобы эти данные можно было копировать, сравнивать и анализировать [13]. Иными словами – это переход от текстовой структуры представления к машиночитываемой структуре. Несколько моделей для экстернализации было уже разработано – SWAN (семантические веб-приложения в нейромедицине), SALT (семантически аннотированный LaTeX) [14], Harmsze и DeWaard [15].

Веб может предоставить перспективные решения для экстернализации. Одной из моделей для публикации семантически богатых научных данных являются нанопубликации [16]. Нанопубликации сериализуются в RDF-формат, и это позволяет использовать их согласно принципам связанных данных. Они могут быть легко опубликованы и агрегированы в Интернете, связаны и использованы повторно.

В целом модель нанопубликаций подходит для идеи Веба научных данных и может быть успешно использована для публикации семантически обогащенных научных данных.

Разработка автоматизированной системы публикации открытых научных данных

На основе анализа описанных систем было принято решение о разработке собственного приложения, которое сможет объединить в себе решение проблемы публикации новых данных в семантически богатых форматах (RDF) и конвертации уже опубликованных результатов. Был разработан прототип системы.

При интеграции данных выявились две проблемы: интеграция с системами, которые не поддерживают семантические форматы, и интеграция с уже существующими наборами данных, которые были созданы согласно принципам связанных данных.

В первом случае необходимо сначала извлечь данные из исходного источника и трансформировать их в семантические форматы. Наиболее распространенным является вариант извлечения данных из реляционных БД, таких как Oracle, MySQL и т.д.

Во втором случае данные уже представлены в машиночитаемом формате, например, в виде RDF. Здесь нет необходимости выполнять их преобразование; вместо этого данные могут быть получены и интегрированы при помощи так называемых федеративных запросов. Такие запросы позволяют осуществлять одновременный поиск различных ресурсов из нескольких распределенных источников данных. В настоящее время разработано несколько федеративных поисковых систем для RDF-данных, которые могут быть разделены на четыре основные категории [17].

1. SPARQL Endpoint federation – системы, которые используют SPARQL-запросы для различных RDF-данных. Преимущество такого подхода заключается в том, что данные получаются на основе исходного набора с возможностью дальнейшей синхронизации. Но использовать такие системы не всегда возможно, так как не все наборы данных, доступные в сети, имеют SPARQL-точки доступа. К примерам таких систем можно отнести HiBISCus, FedX, SPLENDID, ANAPSID, LHD.
2. Linked Data Federation – системы, которые реализуют принципы связанных данных путем поиска по URI во время выполнения. Они не нуждаются в SPARQL-точках доступа, но выполнение запросов может быть более трудоемким. Примерами таких систем являются LDQPS, SIHJoin, WoDQA.
3. Distributed Hash tables – системы используют индексированные хеш-таблицы для федеративных SPARQL-запросов между различными RDF-данными. Для этого требуется RDF-узел для хранения на верхнем уровне распределенной хеш-таблицы, что может существенно снизить нагрузку на сеть. Пример такой системы – ATLAS.
4. Гибридные системы используют SPARQL-точки доступа и поиск по URI для федеративных запросов к RDF-данным. Пример такой системы – ADERIS-Hybrid.

Таким образом, использование семантических технологий за счет представления данных о качестве RDF-модели является наиболее подходящим решением задачи. Такой подход позволит соединить как неструктурированные данные в различных форматах, так и семантически обогащенные данные.

В разработанной системе используется программный продукт Virtuoso Universal Server для извлечения информации из реляционной БД университета. БД университета содержит информацию различных типов: биографические данные сотрудников университета, используемые для публикации на ресурсах университета; данные о публикационной активности университета; данные о научных результатах лабораторий и др.

На первом этапе реализованы доступ к данным Университета ИТМО, таким как публикации сотрудников университета, и публикация данных о них в формате RDF. Принцип работы разработанного прототипа состоит в следующем.

1. Специалист по знаниям разрабатывает принцип (алгоритм) преобразования (отображения) данных из БД университета в набор данных RDF, используя так называемый маппинг-файл, указывает связи между данными.
2. Система проводит выгрузку из БД университета и преобразует их в RDF-формат с использованием программного средства sparqlify.
3. Происходит загрузка данных в Virtuoso Universal Server.

На рис. 1 представлена диаграмма последовательности разработанного прототипа.

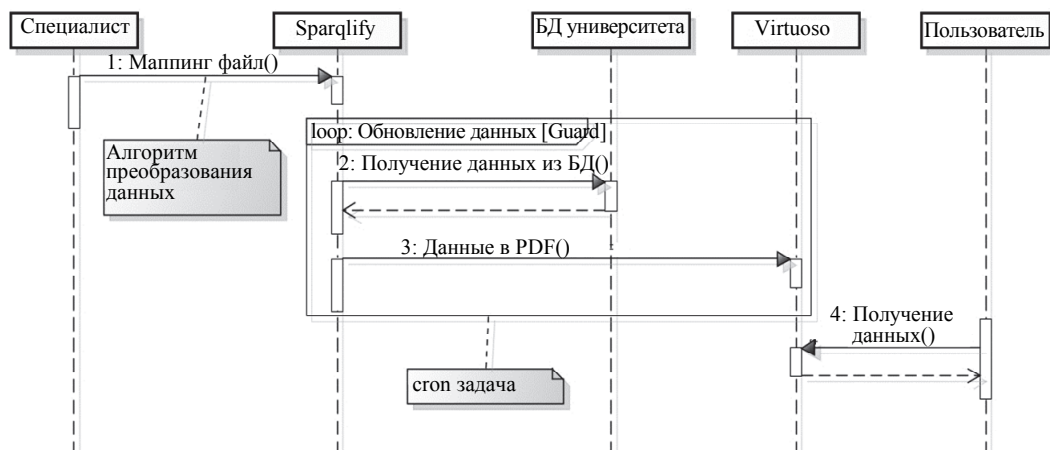


Рис. 1. Диаграмма последовательности прототипа

На рис. 2 представлена диаграмма инфраструктуры для демонстрации хранения данных, хранения маппингов, участие пользователя и т.д.

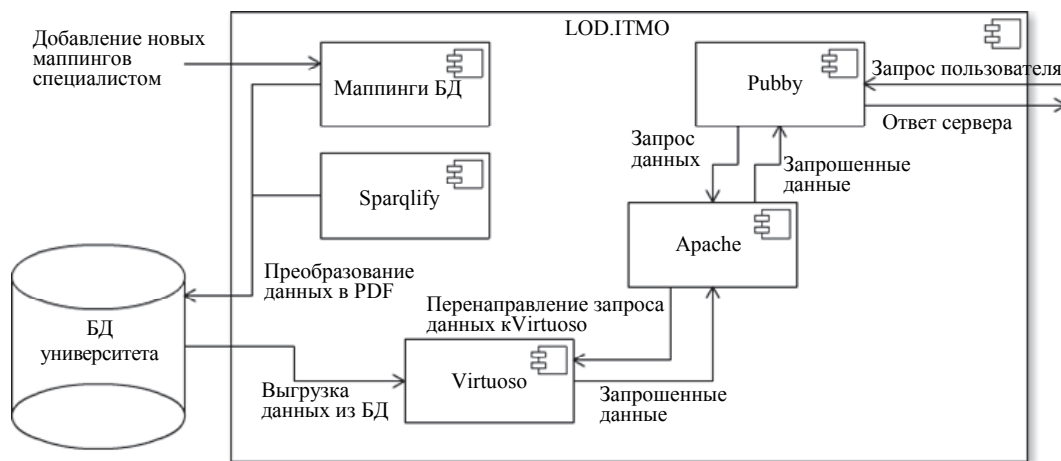


Рис. 2. Диаграмма инфраструктуры

```

PREFIX ifmolod:<http://lod.ifmo.ru/>
PREFIX vivoplus:<http://vivoplus.aksw.org/ontology#>
PREFIX vivo:<http://vivoweb.org/ontology/core#>
PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>
PREFIX foaf:<http://xmlns.com/foaf/0.1/>
PREFIX xsd:<http://www.w3.org/2001/XMLSchema#>

CREATE VIEW Staff AS CONSTRUCT {
  ?person a foaf:Person ;
    vivoplus:pcardId ?pcardId;
    foaf:firstName ?firstName;
    vivo:middleName ?middleName;
    foaf:lastName ?lastName;
    rdfs:label ?label;
    vivo:affiliatedOrganization ?laboratory;
    a ?supervisorType.
}
WITH
  ?person = uri(concat("http://lod.ifmo.ru/Person", ?PCARD_ID))
  ?pcardId = typedLiteral(?PCARD_ID,xsd:positiveInteger)
  ?firstName = plainLiteral(?FIRST_NAME, 'ru')
  ?middleName = plainLiteral(?PATRONYMIC, 'ru')
  ?lastName = plainLiteral(?LAST_NAME, 'ru')
  ?label = plainLiteral(concat(concat(?FIRST_NAME," "),?LAST_NAME))
  ?laboratory = uri(ifmolod:,"Laboratory",?NET_DEP_ID)
  ?supervisorType = uri(vivoplus:;?SUPERVISOR_TYPE)
FROM
  [[SELECT NET_DEP_ID, PCARD_ID,FIRST_NAME, PATRONYMIC, LAST_NAME,
  replace(replace(RANG_BOSS, 'Научный_руководитель', 'ScientificSupervisor'),'Руководитель', 'Supervisor') SUPERVISOR_TYPE
  FROM sem_person_fio]]
    
```

Рис. 3. Пример маппинг-файла

Преобразование данных из БД в RDF реализуется при помощи программного средства sparqlify. Входным файлом для него является создаваемый специалистом маппинг-файл, который и описывает принципы конвертации данных. Структура маппинг-файла похожа на структуру SQL-запроса к базе данных: описывается, какие поля будут добавлены в RDF и какие преобразования данных при этом будут реализованы (например, реализация связи между данными).

Пример маппинга для получения персональных данных пользователя (ФИО, места работы – лаборатории) представлен на рис. 3.

В разработанном прототипе доступны данные (выгруженные с БД университета) о научных сотрудниках, лабораториях, публикациях. Выгруженные данные можно использовать в различных приложениях, таких как, например, поиск.

Заключение

В работе проанализированы существующие тенденции в области публикации открытых научных данных и основные способы их публикации. Выявлены проблемы существующих решений, предложен прототип автоматизированной системы, предназначенной для их решения, на примере которого показана эффективность и новые возможности по поиску научной информации, опубликованной в формате семантических связанных данных. В дальнейшем авторы планируют расширить систему, добавив в нее возможность связывания данных с другими открытыми датасетами для осуществления федеративного поиска.

Литература

1. Keßler C., D'Aquin M., Dietze S. Linked data for science and education // *Semantic Web*. 2013. V. 4. N 1. P. 1–2. doi: 10.3233/SW-120091
2. Larsen P.O., von Ins M. The rate of growth in scientific publication and the decline in coverage provided by Science Citation Index // *Scientometrics*. 2010. V. 84. N 3. P. 575–603. doi: 10.1007/s11192-010-0202-z
3. Das S., Sundara S., Cyganiak R. R2RML: RDB to RDF Mapping Language [Электронный ресурс]. Режим доступа: <http://www.w3.org/TR/r2rml/>, свободный. Яз. англ. (дата обращения 06.05.2015).
4. Sjaaveland M.G., Lian E.H., Horrocks I. Publishing the Norwegian Petroleum Directorate's FactPages as semantic web data // *Lecture Notes in Computer Science*. 2013. V. 8219. N 2. P. 162–177. doi: 10.1007/978-3-642-41338-4_11
5. Rodriguez J.B. et al. R2O, an extensible and semantically based database-to-ontology mapping language // *Proc. 2nd Workshop on Semantic Web and Databases*. 2004. V. 3372. P. 1069–1070.
6. VirtuosoUniversalServer [Электронный ресурс]. Режим доступа: <http://www.w3.org/wiki/VirtuosoUniversalServer>, свободный. Яз. англ. (дата обращения 21.01.2015).
7. Leinberger M., Scheglmann S., Lammel R., Staab S., Thimm M., Viegas E. Semantic web application development with LITEQ // *Lecture Notes in Computer Science*. 2014. V. 8797. P. 212–227.
8. Heath T., Bizer C. *Linked Data: Evolving the Web into a Global Data Space*. 1st ed. Morgan & Claypool Publ., 2011. 136 p. doi: 10.2200/S00334ED1V01Y201102WBE001
9. Microsoft Academic Search [Электронный ресурс]. Режим доступа: <http://academic.research.microsoft.com>, свободный. Яз. англ. (дата обращения: 20.08.2015).
10. Devare M., Corson-Rikert J., Caruso B., Lowe B., Chiang K., McCue J. Connecting people, creating a virtual life sciences community // *D-Lib Magazine*. 2007. V. 13. N 7. P. 1082–9873. doi: 10.1045/july2007-devare
11. Krafft D.B., Cappadona N.A., Caruso B., Corson-Rikert J., Devare M., Lowe B. VIVO: Enabling national networking of scientists // *Proc. Web Science Conference*. Raleigh, USA, 2010. V. 2010. P. 1310–1313.
12. Nonaka I., Takeuchi H. *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*. NY: Oxford University Press, 1995. 304 p.
13. Groza T., Handschuh S., Clark T., Shum S.B., de Waard A. A Short Survey of Discourse Representation Models [Электронный ресурс]. Режим доступа: <http://ceur-ws.org/Vol-523/Groza.pdf>, свободный. Яз. англ. (дата обращения 20.08.2015).
14. Groza T., Handschuh S., Moller K., Decker S. SALT - Semantically annotated LaTeX for scientific publications // *Lecture Notes in Computer Science*. 2007. V. 4519. P. 518–532.
15. de Waard A., Breure L., Kircz J.G., van Oostendorp H. Modeling Rhetoric in Scientific Publications [Электронный ресурс]. Режим доступа: http://www.researchgate.net/publication/46680525_Modeling_Rhetoric_in_Scientific_Publications, свободный. Яз. англ. (дата обращения: 20.08.2015).
16. Sernadela P., van der Horst E., Thompson M., Lopes P., Roos M., Oliveira J.L. A nanopublishing architecture for biomedical data // *Proc. 8th Int. Conf. on Practical Applications of Computational Biology and Bioinformatics*, PACBB. Salamanca, Spain, 2014. V. 294. N 6. P. 277–284. doi: 10.1007/978-3-319-07581-5_33

17. Saleem M., Khan Y., Hasnain A., Ermilov I., Ngonga Ngomo A.-C. A fine-grained evaluation of SPARQL endpoint federation systems // Semantic Web Journal. 2015. V. 6. N 6. doi: 10.3233/SW-150186

- Муромцев Дмитрий Ильич** – кандидат технических наук, доцент, заведующий кафедрой, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, d.muromtsev@gmail.com
- Леманн Йенс** – кандидат технических наук, руководитель исследовательской группы, Лейпцигский Университет, Лейпциг, 04109, Германия, i.semerhanov@gmail.com
- Семерханов Илья Александрович** – кандидат технических наук, научный сотрудник, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, i.semerhanov@gmail.com
- Навроцкий Михаил Александрович** – аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, m.navrotskiy@gmail.com
- Ермилов Иван Сергеевич** – инженер-специалист, научный сотрудник, Лейпцигский Университет, Лейпциг, 04109, Германия, iermilov@informatik.uni-leipzig.de
- Dmitry I. Mouromtsev** – PhD, Associate professor, Head of Chair, ITMO University, Saint Petersburg, 197101, Russian Federation, d.muromtsev@gmail.com
- Jens Lehmann** – PhD, Research Group Leader, University of Leipzig, Leipzig, 04109, Germany, lehmann@informatik.uni-leipzig.de
- Илья А. Semerhanov** – PhD, scientific researcher, ITMO University, Saint Petersburg, 197101, Russian Federation, i.semerhanov@gmail.com
- Mikhail A. Navrotskiy** – postgraduate, ITMO University, Saint Petersburg, 197101, Russian Federation, m.navrotskiy@gmail.com
- Ivan S. Ermilov** – scientific researcher, University of Leipzig, Leipzig, 04109, Germany, iermilov@informatik.uni-leipzig.de