

УДК 004.65, 004.053, 005, 001.5

СЕМАНТИЧЕСКАЯ БИБЛИОТЕКА КАК СРЕДСТВО ОПРЕДЕЛЕНИЯ НАУЧНОЙ ПРЕДМЕТНОЙ ОБЛАСТИ

О. М. Атаева¹ [0000-0003-0367-5575], В. А. Серебряков² [0000-0003-1423-621X]

^{1,2} Вычислительный центр им. А.А. Дородницына ФИЦ ИУ РАН, Москва,
ул. Вавилова, 40

¹ oli@ultimeta.ru

² serebr@ultimeta.ru

Аннотация

Рассмотрены информационная система, предназначенная для представления предметной области, связанной с наукой, и ее особенности. Выделены общие концепции для формального описания такой предметной области в базе знаний семантической библиотеки. Особенность этих областей заключается в том, что структура данных подвержена частым изменениям. Поэтому средство организации знаний, в качестве которого выступает семантическая библиотека, должно быть достаточно универсальным и не требовать глубоких технических познаний. В работе приведены описание функциональности системы и ее использования при настройке на предметную область. Для каждой области набор ресурсов может отличаться как по формату, так и по набору самих ресурсов. Набор понятий, формирующих описание контента библиотеки, должен быть настолько универсальным, чтобы мог адаптироваться под нужды конкретной области. Для представления данных использованы метаданные трех уровней.

Ключевые слова: семантическая библиотека, онтология, представление знаний

ВВЕДЕНИЕ

Вопросами семантической организации знаний занимались различные исследователи с древнейших времен. Библиотеки, специализированные по конкретным областям, используют обычно свои классификаторы для систематизации своих ресурсов. Такой подход обеспечивает более детальный анализ содержания

документов и соотношение смысловых понятий содержимого библиотеки с определенным направлением специализированной области знания.

Накопленные данные стали доступны широкому кругу пользователей через Сеть, функциональность цифровых библиотек становится все разнообразней, удовлетворяя *информационные потребности* пользователей.

В фокусе настоящей работы лежат предметные области, связанные с наукой, и их особенности. Выделены общие концепции для их формальных описаний в базе знаний. Особенность этих областей заключается в том, что структура данных подвержена частым изменениям [1–4]. Основной акцент сделан на представлении обобщенной модели научной предметной области и ее особенностей, реализации в поисковых системах и отличиях от классических подходов к поиску информации в научных массивах данных.

Новые проблемы и вызовы относятся также к представлению знаний в информационной среде для различных областей науки с использованием современных подходов. Для обеспечения потребления научной информации на новом уровне в первую очередь необходим переход к семантически значимому представлению научных знаний, извлекаемых из информации в цифровой среде.

Для представления данных предметной области используют метаданные трех уровней: (1) универсальные понятия без привязки к предметной области, или метаметаданные; (2) понятия для описания конкретной предметной области или метаданные, определения которых задаются в терминах первого уровня; (3) данные прикладной области как таковые, представленные в терминах метаданных второго уровня. На основе этих метаданных настраиваются интерфейсы взаимодействия с пользователями для навигации, редактирования и поиска информации.

Главной задачей создания и описания обобщенного представления научных знаний для некоторой области является помощь экспертам в организации знаний и предоставления доступа к ней [5–9]. При этом средство организации знаний должно быть достаточно универсальным и не требовать глубоких технических познаний.

Была поставлена задача создания такой информационной системы, которая могла бы учитывать все разнообразие различных типов ресурсов научной предметной области, которые могут в ней храниться, и при этом поддерживать ее терминологическое описание. Фактически такая система должна представлять собой конструктор с адаптируемой моделью контента хранимых данных для создания цифровой библиотеки любой направленности. Адаптируемая модель данных позволяет описать произвольную модель данных контента библиотеки в рамках предметной области, фиксированной в терминах тезауруса. Такая информационная система должна учитывать разнообразие типов ресурсов научной предметной области и при этом поддерживать ее терминологическое описание. Основные задачи такой системы – представление контента предметной области в виде онтологии и поддержка интеграции данных из источников. На данный момент реализован и готов к использованию дистрибутив семантической библиотеки. Ниже дано описание основных идей построения модели данных и подсистем, которые представлены в дистрибутиве информационной системы

1. О МОДЕЛИ ДАННЫХ

В информационную модель семантической библиотеки были введены понятия для описания содержимого библиотеки для некоторой предметной области [10–13]. Эти понятия позволяют сконструировать описание любых типов информационных ресурсов для этой области. При этом согласно определению информационных объекты, являющиеся непосредственно содержимым библиотеки, имеют распределенную природу, что означает, что данные могут поступать из различных источников и агрегировать информацию об информационном объекте из различных источников, непосредственно сохраняя данные в самой библиотеке или сохраняя ссылки на идентичные объекты в источниках данных.

Для описания ресурсов, составляющих контент конкретной предметной области, использованы понятия, общие для любой из них, т. е. набор понятий, формирующих описание контента библиотеки, должен быть настолько универсальным, чтобы мог адаптироваться к нуждам конкретной области.

Контент библиотеки тесно связан с тезаурусом, который поддерживает родственные связи различных типов как между самими концептами, так и между концептами и информационными объектами. Это позволяет реализовать гибкий

настраиваемый поиск, результатом которого будет сбалансированный список объектов предметной области. На основе одного и того же тезауруса определяются коллекции ресурсов самых разнообразных типов. Такой подход чрезвычайно полезен для создания отдельных пользовательских коллекций.

Фактически понятия делятся на три категории: первая включает определения понятий контента семантической библиотеки, вторая категория относится к определению понятий, необходимых для поддержки терминов в тезаурусе предметной области, и третья включает определения, необходимые для описания процессов интеграции контента этих ресурсов [14–23]. На основе этих определений описаны основные процессы, такие, как, например, интеграция данных из разных источников, категоризация/классификация, отображение разных моделей данных источников на заданную предметную область, построение классов эквивалентности и т. д.

2. АРХИТЕКТУРА

Рассмотрим формальное описание системы, определяющее ее цели, функции, внешне видимые свойства и интерфейсы. Оно включает также описание компонентов системы и их отношений наряду с принципами, управляющими ее дизайном, функционированием и возможным последующим развитием. Это описание включает программные подсистемы, визуализированные свойства этих подсистем, отношения между подсистемами и ограничения в их использовании. При этом каждая подсистема может состоять из нескольких уровней абстракции, а каждый уровень может иметь свою архитектуру. Ниже приведен список основных подсистем:

- Подсистема описания контента информационной системы;
- Подсистема управления тезаурусом;
- Подсистема автоматизированной обработки и представления данных;
- Подсистема реализации задач интеграции данных;
- Рекомендательная подсистема.

Каждая из этих подсистем отвечает за определенную функциональность и использует свое подмножество понятий из информационной модели.

3. ПОДСИСТЕМА ОПИСАНИЯ КОНТЕНТА

Рассмотрим одну из подсистем, которая определяет основные настройки системы. За универсальность определения контента системы отвечает набор понятий, составляющих информационную модель контента библиотеки Libmeta: *информационный ресурс* и *информационный объект*, которые описывают экземпляры ресурсов. *Информационный ресурс* является основной единицей описания контента библиотеки, а *информационный объект* представляет экземпляры информационных ресурсов. Каждый из них имеет собственный уникальный идентификатор. Фактически семантическое значение *информационного ресурса* является эквивалентным понятию класса *онтологии* с некоторыми ограничениями в его описании. Структура описания информационных объектов определяется понятиями *атрибут* и *набор атрибутов*, которые определяются при описании соответствующего ресурса. Атрибут является элементом описания свойств ресурса, а набор атрибутов определяется как коллекция атрибутов разных видов. Типы атрибутов следующие: *файловый, объектный, числовой, текстовый, строковый*. Помимо определения круга значений атрибута важной характеристикой являются его тип и определение количества его значений. Для описания конкретного информационного ресурса используется понятие *значение атрибута*, которое тесно связано с понятием *атрибут* и является фактически контейнером для хранения конкретных значений *информационного объекта* определенного типа.

Приведенные понятия обеспечивают структурированное описание контента и поддержку его адаптируемости. Такой подход также обеспечивает описание конкретных ресурсов и их объектов в виде RDF-троек и предоставления SPARQL точки доступа для публикации данных в машиночитаемых форматах.

В общем случае конкретная реализация модели контента библиотеки может быть основана на некоторой импортируемой онтологии, классы которой превращаются в ресурсы, свойства могут быть описаны в терминах атрибутов Libmeta, а наборы атрибутов определяют фактически домены свойств онтологий. При построении модели ресурсов библиотеки на основе этой онтологии сохраняются все URI свойств, отношений и классов выбранной онтологии. При необходимости при импортировании выбранной онтологии в систему можно изменить набор понятий, расширив или, наоборот, сократив его средствами системы.

Конечно, такой способ отображения онтологии на понятия системы LibMeta не сохраняет весь возможный перечень ограничений, накладываемых на свойства и классы онтологии изначально, но структурная ее часть сохраняется, что является достаточным для решения задач, определенных в рамках системы.

На рисунке 1 приведены основные понятия, используемые для конструирования описания предметной области в рамках этой подсистемы.

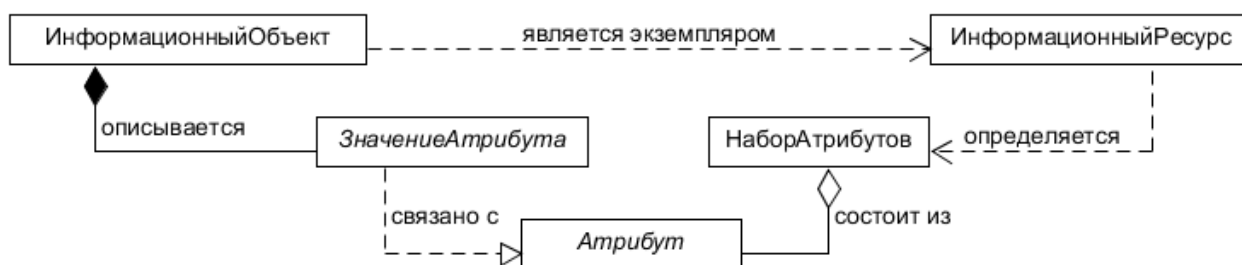


Рис. 1. Основные понятия, используемые для конструирования описания предметной области

При описании *информационных ресурсов* и определении набора их атрибутов важную роль играют *виды атрибутов*, которые формируют структурное описание ресурса. Атрибуты делятся на несколько пересекающихся видов: *поисковые, описательные, административные, идентифицирующие*. При формировании интерфейсов поиска важную роль играют именно *поисковые* атрибуты, которые используются при выполнении атрибутного поиска по типам ресурсов. Результатом такого поиска являются объекты, краткое описание которых представлено пользователю посредством описательных атрибутов.

Фактически в рамках этой подсистемы выполняется первичная настройка конфигурации контента библиотеки и ее интерфейсов под конкретную предметную область. На рисунке 2 изображена последовательность действий пользователя по настройке системы.

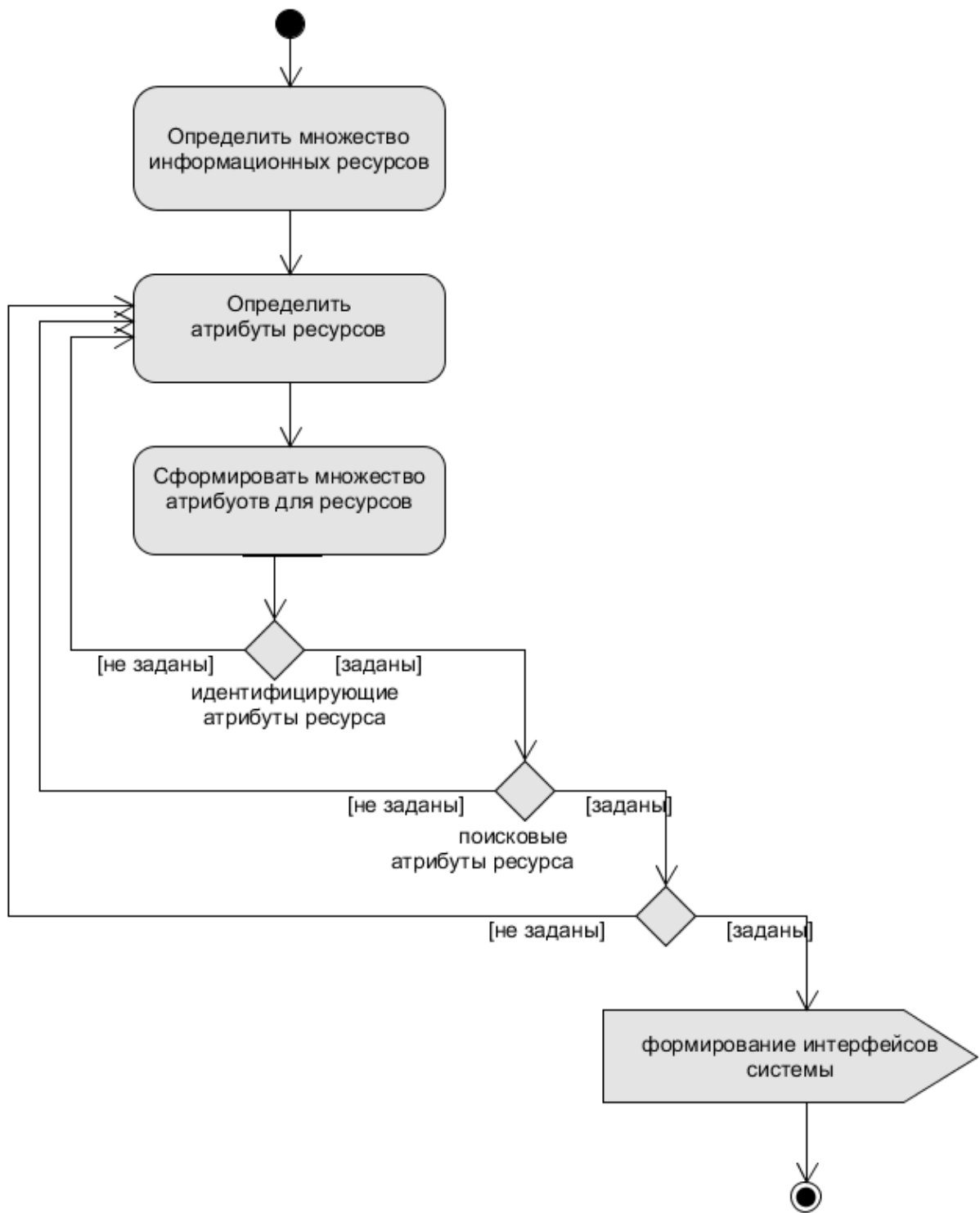


Рис. 2. Последовательность действий пользователя по настройке системы

4. ОСНОВНАЯ ФУНКЦИОНАЛЬНОСТЬ LIVMETA

Эта функциональность такова:

- создание/просмотр/редактирование информационных ресурсов и их структуры;
 - создание/просмотр/редактирование информационных объектов и их структуры;
 - подключение источников данных;
 - загрузка данных из подключенных источников данных, в дальнейшем становящихся частью контента библиотеки;
 - создание/просмотр/редактирование структуры тезауруса поддерживаемой предметной области;
 - создание/просмотр/редактирование понятий тезауруса
 - пакетная загрузка данных, составляющих контент библиотеки;
 - атрибутивный/семантический/полнотекстовый поиск и навигация по доступным информационным объектам системы;
 - атрибутивный/семантический/полнотекстовый поиск по источникам данных;
 - создание/просмотр/редактирование коллекций информационных объектов;
 - формирование онтологии предметной области по описанию структуры информационных ресурсов и тезауруса;
 - предоставление данных, составляющих контент системы в машиночитаемом формате;
 - выделение связей между информационными объектами и понятиями тезауруса;
 - поддержка семантических меток или фолксономии [24 – 25] для описания тематической направленности информационных объектов;
 - создание/просмотр/редактирование области интересов пользователя;
 - создание рекомендательной системы:
 - a. на основе описания интересов пользователя;
 - b. на основе рассматриваемого тезауруса предметной области;
-

- поддержка микротезаурусов пользователей на основе тезауруса предметной области.

Функциональность LibMeta, доступная для всех публичных пользователей:

- просмотр информационных ресурсов и их структуры;
- просмотр информационных объектов и их структуры;
- атрибутивный/семантический/полнотекстовый поиск и навигация по доступным ресурсам системы;
- атрибутивный и семантический поиск по источникам данных;
- просмотр общедоступных коллекций информационных объектов.

С точки зрения авторизированного пользователя, семантическая библиотека дополнительно обеспечивает ему следующую функциональность:

- определение своего микротезауруса как расширение некоторого узла основного терминологического тезауруса, определенного в системе. Также обеспечивается поддержка создания так называемых *аннотационных онтологий* или *онтологий пользователей (фолксономии)*, которые представляют собой коллективный словарь пользователей, составленный в результате процесса проставления ими семантических меток ресурсов;

- определение собственных коллекций информационных объектов;
- организация совместных тематических коллекций для групп пользователей;
- атрибутивный и семантический поиск по источникам данных с возможностью сохранения результатов поиска;

- пользователь в роли администратора системы имеет доступ ко всей вышеопределенной функциональности и может воспользоваться дополнительной, доступной только ему функциональностью:

- a. может по запросу пользователей расширять описания типов ресурсов или создавать новые;
- b. может по запросу пользователей включать их объекты ресурсов в общедоступный список объектов;
- c. для групп пользователей доступны возможности редактирования определенных типов ресурсов или таксономий;

- d. редактировать группы и роли пользователей и набор доступных им операций;
- e. осуществлять редактирование и настройку основного терминологического тезауруса и его связей.

ЗАКЛЮЧЕНИЕ

Представлено описание информационной системы для реализации функциональности семантической библиотеки для некоторой предметной области. В результате эксперты предметной области получают возможность реализации главной задачи библиотеки – *семантического/интеллектуального* конструирования научного пространства знаний для некоторой предметной области, т. е. наделение его семантикой за счет явного выделения интеллектуально значимых связей, поддержки семантической разметки. Основным инструментом конструирования является онтология предметной области, которая позволяет осмысленно структурировать и обеспечить связность между ресурсами, которые включены в научное пространство знаний предметной области, и использование унифицированной терминологической поддержки в виде тезауруса этой предметной области. Для реализации функций открытости научного пространства знаний реализованы возможности интеграции других источников данных и связывания с их данными. Предоставление функциональности для совместной работы над развитием пространства научного знания повышает эффективность проводимых в нем исследований и расширяет возможности по его поддержке в актуальном состоянии.

СПИСОК ЛИТЕРАТУРЫ

1. *Леонова Ю.В., Федотов А.М.* Создание прототипа системы управления информационными ресурсами // Вестник Восточно-Казахстанского гос. техн. университета и журнала «Вычислительные технологии» ИВТ СО РАН. CITech-2018, Усть-Каменогорск, Казахстан. 2018. С. 47–56.

URL: http://www.ict.nsc.ru/jspui/bitstream/ICT/1879/8/Part1_46-55.pdf

2. *Кулагин М.В., Лопатенко А.С.* Научные информационные системы и электронные библиотеки. Потребность в интеграции // Электронные библиотеки: перспективные методы и технологии, электронные коллекции. Сборник докладов

Третьей Всероссийской конференции, Петрозаводск, 11–13 сентября, 2001. С. 14–19. URL: <http://elib.ict.nsc.ru/jspui/handle/ICT/1864>

3. Шокин Ю.И., Федотов А.М., Барахнин В.Б. Проблемы поиска информации. Новосибирск: Наука, 2010. URL: <https://lib.nsu.ru/xmlui/handle/nsu/161>

4. Börner K. et al. VIVO: A semantic approach to scholarly networking and discovery // Synthesis lectures on the Semantic Web: theory and technology. 2012. Vol. 7. No. 1. P. 1–178. <https://doi.org/10.2200/S00428ED1V01Y201207WBE002>

5. Нзюк Н.Б., Тузовский А.Ф. Обзор подходов семантического поиска // Доклады Томского государственного университета систем управления и радиоэлектроники. 2010. № 2-2 (22). С. 234–238. URL: <https://journal.tusur.ru/ru/arhiv/2-2-2010/obzor-podhodov-semanticheskogo-poiska>

6. Апанович Э.В., Винокуров П.С., Кислицина Т.А. Средства визуального анализа информационного наполнения порталов, входящих в облако Linked Open Data // Труды XIII Всерос. науч. конф. «Электронные библиотеки: перспективные методы и технологии, электронные коллекции», г. Воронеж, 14–17 октября, 2011. С. 113–120. URL: <http://ceur-ws.org/Vol-803/paper15.pdf>

7. Оробинская Е.А., Дорошенко А.Ю. Использование онтологий для автоматической обработки текстов на естественном языке // Вестник Нац. техн. ун-та ХПИ: сб. науч. тр. Темат. вып. «Актуальные проблемы развития украинского общества». Харьков: НТУ ХПИ. 2011. № 30. С. 101–106.

URL: <http://repository.kpi.kharkov.ua/handle/KhPI-Press/14950>

8. Добров Б.В., Лукашевич Н.В. Тезаурус Рутез как ресурс для решения задач информационного поиска // Труды Всероссийской конференции Знания–Онтологии–Теории (ЗОНТ-09), Новосибирск. 2009. Т. 10. С. 250–259

URL: <http://ns.math.nsc.ru/conference/zont09/reports/93Dobrov-Lukashevich.pdf>

9. Ngonga Ngomo A.C. et al. Sorry, i don't speak SPARQL: translating SPARQL queries into natural language // Proceedings of the 22nd International Conference on World Wide Web. ACM, 2013. P. 977–988.

URL: <https://dl.acm.org/doi/10.1145/2488388.2488473>

10. Серебряков В.А., Атаева О.М. Основные понятия формальной модели семантических библиотек и формализация процессов интеграции в ней // Программные продукты и системы. 2015. № 4. С. 180–187.

<https://doi.org/10.15827/0236-235x.112.180-187>

11. *Атаева О.М., Серебряков В.А.* Персональная открытая семантическая цифровая библиотека LibMeta. Конструирование контента. Интеграция с источниками LOD // Информатика и её применения. 2017. Т. 11. №2. С. 85–100.

<https://doi.org/10.14357/19922264170210>

12. *Атаева О.М.* Информационная модель семантической библиотеки LibMeta // Программные продукты и системы. 2016. № 4. С. 36–44.

<http://dx.doi.org/10.15827/0236-235X.116.036-044>

13. *Атаева О.М., Серебряков В.А.* Онтология цифровой семантической библиотеки LibMeta // Информатика и её применения. 2018. Т. 12. №1. С. 2–10.

<https://doi.org/10.14357/19922264180101>

14. *Ломов П.А., Шишаев М.Г.* Интеграция онтологий с использованием тезауруса для осуществления семантического поиска // Информационные технологии и вычислительные системы. 2009. № 3. С. 49–59.

URL: <http://mi.mathnet.ru/itvs460>

15. *Katsis Y., Papakonstantinou Y.* View-based data integration // Encyclopedia of Database Systems. 2009. P. 3332–3339.

https://doi.org/10.1007/978-1-4614-8265-9_1072

16. *Xu L., Embley D.W.* Combining the Best of Global-as-View and Local-as-View for Data Integration // ISTA. 2004. Vol. 48. P. 123–136.

URL: <https://subs.emis.de/LNI/Proceedings/Proceedings48/GI.Band.48-9.pdf>

17. *Когаловский М.Р.* Методы интеграции данных в информационных системах. 2010. URL: http://www.ipr-ras.ru/old_site/articles/kogalov10-05.pdf

18. *Карабач А.Е.* Системы интеграции информации на основе семантических технологий // Наука, техника и образование. 2014. № 2 (2).

URL: <https://cyberleninka.ru/article/n/sistemy-integratsii-informatsii-na-osnove-semanticheskikh-tehnologiy>

19. *Lenzerini M.* Data integration: A theoretical perspective // Proceedings of the twenty-first ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems. ACM, 2002. P. 233–246. <http://dx.doi.org/10.1145/543613.543644>

20. *Calvanese D., De Giacomo G., Lenzerini M.* Ontology of Integration and Integration of Ontologies // Description Logics. 2001. Vol. 49. No. 10-19. P. 30.

URL: <http://www.diag.uniroma1.it/degiacom/papers/2001/CaDL01dl.pdf>

21. *Noy N.F.* Semantic integration: a survey of ontology-based approaches // ACM Sigmod Record. 2004. Vol. 33. No. 4. P. 65–70.

<http://dx.doi.org/10.1145/1041410.1041421>

22. *Zhao L., Ichise R.* Ontology integration for linked data // Journal on Data Semantics. 2014. Vol. 3. No. 4. P. 237–254.

<https://doi.org/10.1007/s13740-014-0041-9>

23. *Ле Хоай, Тузовский А.Ф.* Разработка семантических электронных библиотек на основе онтологических моделей // Труды XV Всерос. науч. конф. «Электронные библиотеки: перспективные методы и технологии, электронные коллекции», г. Ярославль, 14–17 октября, 2013. С. 143–151.

URL: <http://ceur-ws.org/Vol-1108/paper18.pdf>

24. *Noruzi A.* Folksonomies:(un) controlled vocabulary? // KO KNOWLEDGE ORGANIZATION. 2006. Vol. 33. No. 4. P. 199–203.

URL: <https://core.ac.uk/download/pdf/11882508.pdf>

25. *Gruber T.* Ontology of folksonomy: A mash-up of apples and oranges // International Journal on Semantic Web and Information Systems (IJSWIS). 2007. Vol. 3. No. 1. P. 1–11. URL: <https://tomgruber.org/writing/ontology-of-folksonomy.htm>

SEMANTIC LIBRARY AS A TOOL OF DEFINING A SCIENTIFIC SUBJECT AREA

O. M. Ataeva¹ [0000-0003-0367-5575], V. A. Serebriakov² [0000-0003-1423-621X]

^{1, 2} *Dorodnicyn Computing Center FRC CSC of RAS*

¹ oli@ultimeta.ru

² serebr@ultimeta.ru

Abstract

The paper considers an information system designed to represent a subject area related to science and its features. Highlighted general concepts for formal descriptions of such a subject area in the knowledge base of the semantic library. The peculiarity of these areas is that the data structure is subject to frequent changes. Therefore, the means of organizing knowledge, which is a semantic library, should be sufficiently universal and not require deep technical knowledge. The paper describes the functionality of the system and its use. For each area, the set of resources can differ both in format and in the set of the resources themselves. The set of concepts that form the description of the library's content should be so universal that it can be adapted to the needs of a particular area. Three levels of metadata are used to represent the data.

Keywords: *semantic library, ontology, knowledge representation*

REFERENCES

1. *Leonova Yu.V., Fedotov A.M.* Sozdanie prototipa sistemy upravleniya informacionnymi resursami //Vestnik Vostochno-Kazahstanskogo gos. Tekhn. Universiteta i zhurnala Vychislitel'nye tekhnologii IVT SO RAN.–CITech-2018, Ust'-Kamenogorsk, Kazahstan. 2018. S. 47–56.

URL: http://www.ict.nsc.ru/jspui/bitstream/ICT/1879/8/Part1_46-55.pdf

2. *Kulagin M.V., Lopatenko A.S.* Nauchnye informacionnye sistemy i elektronnye biblioteki. Potrebnost' v integracii // Digital Libraries: Advanced Methods and Technologies, Digital Collections, Petrozavodsk, September 11–13, 2001. S. 14–19.

URL: <http://elib.ict.nsc.ru/jspui/handle/ICT/1864>

3. *Shokin Yu.I., Fedotov A.M., Barahnin V.B.* Problemy poiska informacii. Novosibirsk: Nauka, 2010. URL: <https://lib.nsu.ru/xmlui/handle/nsu/161>

4. *Börner K. et al.* VIVO: A semantic approach to scholarly networking and discovery // *Synthesis lectures on the Semantic Web: theory and technology*. 2012. Vol. 7. No. 1. P. 1–178. <https://doi.org/10.2200/S00428ED1V01Y201207WBE002>

5. *Ngok N.B., Tuzovskij A.F.* Obzor podhodov semanticheskogo poiska // *Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniya i radioelektroniki*. 2010. № 2-2 (22). S. 234–238.

URL: <https://journal.tusur.ru/ru/arhiv/2-2-2010/obzor-podhodov-semanticheskogo-poiska>

6. *Apanovich Z.V., Vinokurov P.S., Kislicina T.A.* Sredstva vizual'nogo analiza informacionnogo napolneniya portalov, vkhodyashchih v oblako Linked Open Data // *Proceedings of the 13th All-Russian Scientific Conference "Digital libraries: Advanced Methods and Technologies, Digital Collections"* Voronezh, Russia, October 19–22, 2011. S. 113–120. URL: <http://ceur-ws.org/Vol-803/paper15.pdf>

7. *Orobinskaya E.A., Doroshenko A.Yu.* Ispol'zovanie ontologij dlya avtomaticheskoy obrabotki tekstov na estestvennom yazyke // *Vestnik Nac. tekhn. un-ta HPI: sb. nauch. tr. Temat. vyp. "Aktual'nye problemy razvitiya ukrainskogo obshchestva"*. Har'kov: NTU HPI. 2011. № 30. S. 101–106.

URL: <http://repository.kpi.kharkov.ua/handle/KhPI-Press/14950>

8. *Dobrov B.V., Lukashevich N.V.* Tezaurus RuTez kak resurs dlya resheniya zadach informacionnogo poiska // *Trudy Vserossijskoj Konferencii Znaniya-Ontologii-Teorii (ZONT-09)*, Novosibirsk. 2009. T. 10. S. 250–259

URL: <http://ns.math.nsc.ru/conference/zont09/reports/93Dobrov-Lukashevich.pdf>

9. *Ngonga Ngomo A.C. et al.* Sorry, i don't speak SPARQL: translating SPARQL queries into natural language // *Proceedings of the 22nd international conference on World Wide Web*. ACM, 2013. P. 977–988.

URL: <https://dl.acm.org/doi/10.1145/2488388.2488473>

10. *Serebryakov V.A., Ataeva O.M.* Osnovnye ponyatiya formal'noj modeli semanticheskikh bibliotek i formalizaciya processov integracii v nej // *Programmnye produkty i sistemy*. 2015. № 4. S. 180–187.

<https://doi.org/10.15827/0236-235x.112.180-187>

11. *Ataeva O.M., Serebryakov V.A.* Personal'naya otkrytaya semanticheskaya cifrovaya biblioteka LibMeta. Konstruirovaniye kontenta. Integraciya s istochnikami LOD // *Informatika i eyo primeneniya*. 2017. T. 11, No. 2. S. 85–100.

<https://doi.org/10.14357/19922264170210>

12. *Ataeva O.M.* Informacionnaya model' semanticheskoy biblioteki LibMeta // *Programmnye produkty i sistemy*. 2016. № 4. S. 36–44.

<http://dx.doi.org/10.15827/0236-235X.116.036-044>

13. *Ataeva O.M., Serebryakov V.A.* Ontologiya cifrovoj semanticheskoy biblioteki LibMeta // *Informatika i eyo primeneniya*. 2018. T. 12, No. 1. S. 2–10.

<https://doi.org/10.14357/19922264180101>

14. *Lomov P.A., Shishaev M.G.* Integraciya ontologij s ispol'zovaniem tezaurusa dlya osushchestvleniya semanticheskogo poiska // *Informacionnye tekhnologii i vychislitel'nye sistemy*. 2009. № 3. S. 49–59. URL: <http://mi.mathnet.ru/itvs460>

15. *Katsis Y., Papakonstantinou Y.* View-based data integration // *Encyclopedia of Database Systems*. 2009. P. 3332–3339.

https://doi.org/10.1007/978-1-4614-8265-9_1072

16. *Xu L., Embley D.W.* Combining the Best of Global-as-View and Local-as-View for Data Integration // *ISTA*. 2004. Vol. 48. P. 123–136.

URL: <https://subs.emis.de/LNI/Proceedings/Proceedings48/GI.Band.48-9.pdf>

17. *Kogalovskij M.R.* Metody integracii dannyh v informacionnyh sistemah. 2010.

URL: http://www.ipr-ras.ru/old_site/articles/kogalov10-05.pdf

18. *Karabach A.E.* Sistemy integracii informacii na osnove semanticheskikh tekhnologij // *Nauka, tekhnika i obrazovanie*. 2014. № 2 (2).

URL: <https://cyberleninka.ru/article/n/sistemy-integratsii-informatsii-na-osnove-semanticheskikh-tehnologiy>

19. *Lenzerini M.* Data integration: A theoretical perspective // *Proceedings of the twenty-first ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*. ACM, 2002. P. 233–246. <http://dx.doi.org/10.1145/543613.543644>

20. *Calvanese D., De Giacomo G., Lenzerini M.* Ontology of Integration and Integration of Ontologies // *Description Logics*. 2001. Vol. 49. No. 10-19. P. 30.

URL: <http://www.diag.uniroma1.it/degiacom/papers/2001/CaDL01dl.pdf>

21. *Noy N.F.* Semantic integration: a survey of ontology-based approaches // ACM Sigmod Record. 2004. Vol. 33. No. 4. P. 65–70.

<http://dx.doi.org/10.1145/1041410.1041421>

22. *Zhao L., Ichise R.* Ontology integration for linked data // Journal on Data Semantics. 2014. Vol. 3. No. 4. P. 237–254.

<https://doi.org/10.1007/s13740-014-0041-9>

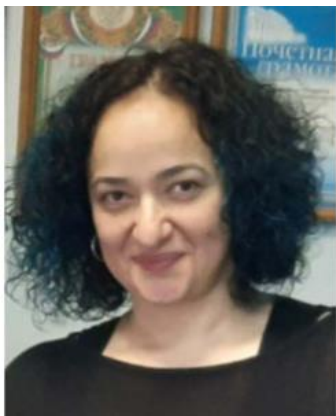
23. *Le Hoaj, Tuzovskij A.F.* Razrabotka semanticheskikh elektronnyh bibliotek na osnove ontologicheskikh modelej // 15th All-Russian Scientific Conference “Digital Libraries: Advanced Methods and Technologies, Digital Collections”. Yaroslavl, Russia, October 14–17, 2013. S. 143–151. URL: <http://ceur-ws.org/Vol-1108/paper18.pdf>

24. *Noruzi A.* Folksonomies:(un) controlled vocabulary? // KO KNOWLEDGE ORGANIZATION. 2006. Vol. 33. No. 4. P. 199–203.

URL: <https://core.ac.uk/download/pdf/11882508.pdf>

25. *Gruber T.* Ontology of folksonomy: A mash-up of apples and oranges // International Journal on Semantic Web and Information Systems (IJSWIS). 2007. Vol. 3. No. 1. P. 1–11. URL: <https://tomgruber.org/writing/ontology-of-folksonomy.htm>

СВЕДЕНИЯ ОБ АВТОРАХ



АТАЕВА Ольга Муратовна – научный сотрудник Вычислительного центра им. А.А. Дородницына ФИЦ ИУ РАН, кандидат техн. наук, специалист в области системного программирования и баз данных.

Olga Muratovna ATAeva – researcher of the of Dorodnicyn computing center FRC SCS RAS, PhD

email: oli@ultimeta.ru

ORCID: 0000-0003-0367-5575



СЕРЕБРЯКОВ Владимир Алексеевич – специалист в области теории формальных языков и её приложений, доктор физ.-мат. наук, профессор, зав. Отделом Вычислительного центра им. А.А. Дородницына ФИЦ ИУ РАН. Руководитель и участник разработки ряда известных программных проектов, в частности, ИСИР и ИСИР РАН, «Научный портал РАН».

Vladimir Alekseevich SEREBRIAKOV – expert in the field of theory of formal languages and its applications, doctor of sciences, professor, head of Dorodnicyn computing center FRC SCS RAS department. Head and participant in the development of a number of well-known program projects, in particular, ISIR and ISIR RAS, Scientific portal RAS.

email: serebr@ultimeta.ru

ORCID: 0000-0003-1423-621X

Материал поступил в редакцию 2 ноября 2021 года