

2015 года, Россия): Труды конференции/ под ред. Л.А. Калиниченко, С.О. Старкова. – Обнинск: ИАЕЭ НИЯУ МИФИ, 2015. 525 с.

4. Standard ECMA-376: Office Open XML File Formats. URL: <http://www.ecmainternational.org/publications/standards/Ecma-376.htm>

5. Хайдаров Ш.М. Методы управления математическим контентом в информационных издательских системах // Тр. Матем. центра им. Н.И. Лобачевского. Материалы 14-й Всерос. Молодежной школы-конференции «Лобачевские чтения–2015 (Казань, 22–27 октября 2015 года). Казань. 2015. С. 162–165.

6. Хайдаров Ш.М. Семантический анализ документов в системе управления цифровыми научными коллекциями // Электронные библиотеки. 2015. Т. 18. № 1–2. С. 61–85.

УДК 004.91

ЭКОСИСТЕМА ONTOMATH И ПРОЕКТ ВСЕМИРНОЙ ЦИФРОВОЙ МАТЕМАТИЧЕСКОЙ БИБЛИОТЕКИ

**А.М. Елизаров¹, Н.Г. Жильцов², А.В. Кириллович³,
Е.К. Липачёв⁴, О.А. Невзорова⁵,**

Казанский (Приволжский) федеральный университет

1 – amelizarov@gmail.com,

2 – nikita.zhiltsov@gmail.com, 3 – alik.kirillovich@gmail.com,

4 – elipachev@gmail.com, 5 – onevzoro@gmail.com

Описаны возможности использования при проведении новых исследований всего корпуса накопленных научных знаний. Такое использование предполагает повсеместное внедрение информационно-коммуникационных технологий (ИКТ), обеспечивающих оптимальное управление имеющимися знаниями, организацию эффективного доступа к ним, а также совместное и многократное использование новых видов структур знаний. Наибольший эффект от внедрения современных ИКТ для дальнейшей организации научных знаний и повышения их понятности можно ожидать в области математики. Эти ожидания в полной мере подтверждены проектом создания Всемирной цифровой математической библиотеки (World Digital Mathematical Library – WDML). Представлены основные направления реализации проекта WDML и результаты по созданию экосистемы OntoMath как его составной части.

Ключевые слова: *WDML, Всемирная цифровая математическая библиотека, экосистема OntoMath, онтологии, семантический поиск*

В настоящее время благодаря повсеместному внедрению информационно-коммуникационных технологий (ИКТ) в научно-исследовательскую деятельность стало возможным при проведении новых

исследований использовать весь корпус накопленных научных знаний. Такое использование предполагает создание комплекса технологий, обеспечивающих оптимальное управление имеющимися знаниями, организацию эффективного доступа к ним, а также совместное и многократное использование новых видов структур знаний. По-видимому, наибольший эффект от внедрения современных ИКТ для дальнейшей организации научных знаний и повышения их понятности можно ожидать в области математики, которая характеризуется наибольшей по сравнению с другими науками формализацией результатов. Эти ожидания в полной мере подтверждены проектом создания Всемирной цифровой математической библиотеки (World Digital Mathematical Library – WDML). Термин WDML введен в 2006 году на Генеральной ассамблее международного математического союза (см. [1]).

Назначение WDML – объединить в распределенной системе взаимосвязанных хранилищ оцифрованные версии всего корпуса математической научной литературы, включая как современные источники, так и ставшие уже историческими (см. [2, 3]). Основные задачи построения WDML и технологии, необходимые для их решения, описаны в [4]. В частности, проект предполагает, что следующим шагом в продвижении математики будут выход за пределы традиционных математических публикаций и построение сети информации, основанной на знаниях, содержащихся в этих публикациях. Благодаря сочетанию методов машинного обучения и усилий редакций и редколлегий математических научных журналов, значительная часть информации и знаний (как связанных открытых данных) в глобальном математическом корпусе станет доступной для исследователей через WDML. Частью проекта WDML, связанной с семантическим представлением математического знания, стал симпозиум, прошедший в 2016 году в Филдсовском институте, г. Торонто [5]. В работе симпозиума приняло участие 37 приглашенных экспертов. В нашем докладе [6] на симпозиуме представлены технологии управления математическими знаниями на основе онтологий (см. также [7–12]). Результатом исследований, описанных в этих работах, стало создание основ цифровой экосистемы OntoMath, основные положения которой полностью коррелируют с идеологией проекта WDML.

OntoMath – это экосистема онтологий, инструментов текстовой аналитики и приложений для управления математическими знаниями. Кратко опишем ее основные элементы.

Платформа семантической публикации – это центральный элемент экосистемы (см. [7]). Эта платформа принимает на вход коллекцию математических публикаций в формате LaTeX и строит их

семантическое представление, интегрированное в облако Linked Open Data. Семантическое представление публикаций включает:

- метаданные: названия, даты, авторы, аффилиации, наименования журналов и т.д.; метаданные описаны с использованием онтологии АКТ Portal;
- логическую структуру публикаций: раздел, теорема, доказательство, формула и т. д., которая описана с помощью онтологии Mocassin;
- терминологию, выраженную посредством онтологии OntoMathPro;
- формулы: переменные внутри них привязаны к понятиям, которые эти переменные обозначают.

Mocassin – это онтология логической структуры математических документов, предназначенная для автоматического анализа математических публикаций в формате LaTeX. Онтология описывает семантику структурных элементов математических документов (например, теоремы, леммы, доказательства, определения и т. д.) и связей между ними (см. [8]).

OntoMathPro – это онтология математического знания, которая организована в виде двух иерархий:

- иерархии областей математики: математическая логика, теория множеств, алгебра, геометрия, топология и т. д.;
- иерархии математических объектов: множество, функция, интеграл, элементарное событие, многочлен Лагранжа и т. д.).

OntoMathPro содержит пять типов отношений: Класс \rightarrow Подкласс, Определяется с помощью, Ассоциативная связь, Задача \rightarrow Метод решения и Область математики \rightarrow Математический объект. Концепты онтологии содержат название концепта на русском и английском языках, определение, ссылки на внешние ресурсы из облака Linked Open Data и связи с другими концептами (см. [9–11]).

Семантический поисковик по математическим формулам: отыскивает формулы, содержащие переменную, обозначающую заданное математическое понятие (например, формулы, содержащие переменную, обозначающую угол, или формулы, связывающие давление и массу).

Рекомендательная система математических публикаций: для заданной публикации строит список «похожих» статей (см. [12]).

Благодарности. Работа выполнена при финансовой поддержке РФФИ (проекты №№ 15-07-08522, 15-47-02472).

Литература

1. Digital Mathematics Library: a vision for the future. International Mathematical Union, 2006, URL: http://www.mathunion.org/fileadmin/IMU/Report/dml_vision.pdf.
2. Olver P.J. What's happening with the World Digital Mathematics Library? URL: http://www.math.umn.edu/~olver/t/_wdmlb.pdf.
3. Olver P.J. The World Digital Mathematics Library: report of a panel discussion // Proceedings of the International Congress of Mathematicians, August 13–21, 2014, Seoul, Korea. Kyung Moon SA, 2014. V. 1. P. 773–785.
4. Developing a 21st century global library for mathematics research. Washington, D.C.: The National Academies Press, 2014. 131 p.
5. Semantic representation of mathematical knowledge workshop, 5 February 2016. URL: <https://www.fields.utoronto.ca/programs/scientific/15-16/semantic/>
6. Elizarov A.M., Zhiltsov N.G., Kirillovich A.V., Lipachev E.K., Nevzorova O.A., Solovyev V.D. The OntoMath ecosystem: ontologies and applications for math knowledge management // Semantic Representation of Mathematical Knowledge Workshop 5 February 2016. URL: <http://www.fields.utoronto.ca/video-archive/2016/02/2053-14698>.
7. Nevzorova O., Zhiltsov N., Zaikin D., Zhibrik O., Kirillovich A., Nevzorov V., Birialtsev E. Bringing Math to LOD: a semantic publishing platform prototype for scientific collections in mathematics // 12th International Semantic Web Conference, Sydney, NSW, Australia, October 21–25, 2013, Proceedings, Part I. Lecture Notes in Computer Science, Vol. 8218. Springer Berlin Heidelberg, 2013. P. 379–394.
8. Solovyev V., Zhiltsov N. Logical structure analysis of scientific publications in mathematics // In: Proceedings of the International Conference on Web Intelligence, Mining and Semantics (WIMS'11). ACM. 2011. P. 21:1–21:9.
9. Nevzorova O., Zhiltsov N., Kirillovich A., Lipachev E. OntoMathPro ontology: a linked data hub for mathe-matics // Communications in Computer and Information Science. 2014. V. 468. P. 105–119.
10. Elizarov A., Kirillovich A., Lipachev E., Nevzorova O., Solovyev V., Zhiltsov N. Mathematical knowledge representation: semantic models and formalisms // Lobachevskii Journal of Mathematics, 2014. V. 35, No 4. P. 347–353.
11. Elizarov A.M., Lipachev E.K., Nevzorova O.A., Solov'ev V.D. Methods and means for semantic structuring of electronic mathematical documents // Dokl. Math. 2014. V. 90, No 1. P. 521–524.
12. Елизаров А.М., Жижченко А.Б., Жильцов Н.Г., Кириллович А.В., Липачёв Е.К. Онтологии математического знания и рекомендательная система для коллекций физико-математических документов // Докл. РАН. 2016. Т. 467, №4. С. 392–395.