

УДК 004.4
ББК 30ф

ЕЛИЗАРОВ А.М.¹, ЛИПАЧЁВ Е.К.², ХАЙДАРОВ Ш.М.³

Институт математики и механики им. Н.И. Лобачевского
Казанского (Приволжского) федерального университета
Казань, Россия

¹ amelizarov@gmail.com, ² elipachev@gmail.com, ³ 15jkeee@gmail.com

СТРУКТУРА И СЕРВИСЫ ЦИФРОВОЙ МАТЕМАТИЧЕСКОЙ БИБЛИОТЕКИ LOBACHEVSKII-DML

***Аннотация:** Представлены результаты разработки электронной математической библиотеки Lobachevskii-DML. Описаны цели создания этой электронной библиотеки, методы управления математическим контентом на основе семантических технологий, взаимодействие с информационными системами научных журналов, система сервисов по поддержке жизненного цикла математического документа.*

***Ключевые слова:** электронные публикации, интеграция данных, электронные библиотеки, электронные математические библиотеки, системы управления информацией, электронная библиотека Lobachevskii DML.*

ELIZAROV A.M.¹, LIPACHEV E.K.², KHAYDAROV S.M.³

N.I. Lobachevskii Institute of Mathematics and Mechanics,
Kazan (Volga Region) Federal University
Kazan, Russia

¹ amelizarov@gmail.com, ² elipachev@gmail.com, ³ 15jkeee@gmail.com

STRUCTURE AND SERVICES DIGITAL MATHEMATICAL LIBRARY LOBACHEVSKII-DML

***Abstract:** We present the results of the development of the electronic mathematical library Lobachevskii-DML. We describe the goals of this electronic library, the methods of managing mathematical content on the basis of semantic technologies. We note the interaction with the information systems of scientific journals. We are a system of services to support the life cycle of a mathematical document.*

***Keywords:** electronic publishing, library automation, machine-actionable digital library, digital mathematics library, DML, WDML, Lobachevskii-DML.*

Введение

Перевод знаний в цифровую форму и перемещение научных коммуникаций в сетевое пространство не только изменили сложившуюся научную инфраструктуру, но и поставили новые задачи по управлению знаниями. Как известно, необходимым элементом научного исследования является описание связей новых научных результатов с ранее полученными. В современных условиях для выполнения этого требования необходимо наличие в сети научного контента, как современного, так и ставшего уже классическим. В определенной степени задачу перехода от хранения бумажных документов к управлению цифровым контентом решают традиционные библиотеки, использующие в своей работе сетевой инструментарий и тем самым существенно расширившие стандартные функции библиотек. Частью формируемой новой научной электронной инфраструктуры являются электронные библиотеки (см., например, [1, 2]).

Особенности математических документов (а это не только преобладание формульного содержания над текстовым, но и логическая структура документов, состоящих из определений, теорем и доказательств) требуют специальных методов управления электронным контентом. Это принципиально важно даже для таких стандартных функций, реализуемых в Сети, как поиск, аннотирование и формирование рекомендаций (см., например, [3–6]).

С 2001 года в рамках различных международных проектов разрабатываются методы управления электронным математическим контентом. Основной целью этих проектов является создание общей инфраструктуры электронных математических библиотек, обеспечивающей доступ к математическим знаниям, сконцентрированным в локальных коллекциях научных журналов, сборников и книг. Например, Европейский проект EuDML (<https://eudml.org/>) направлен на интеграцию европейских математических ресурсов (см., например, [7]). Глобальная инициатива WDML (World Digital Mathematical Library) ставит основной задачей объединение в распределенной системе электронных коллекций всего корпуса электронных математических документов [8]. Определяющая роль в этом объединении отведена национальным электронным математическим библиотекам. Примерами таких библиотек являются Общероссийский математический портал Math-Net.Ru [9], Czech Digital Mathematics Library (DML-CZ, <http://dml.cz/>) [10], Centre de Diffusion de Revues Académiques Mathématiques (CEDRAM, <http://www.cedram.org/>), Numerisation de Documents Anciens Mathématiques (NUMDAM, <http://www.cedram.org/>).

В настоящей работе представлена электронная библиотека Lobachevskii Digital Mathematical Library (Lobachevskii-DML, <http://www.lobachevskii-dml.ru/>), объединяющая электронные коллекции математических документов Казанского университета и предоставляющая возможности навигации по понятиям и объектам, выделенным из документов.

Структура Lobachevskii-DML

Lobachevskii-DML — это цифровая математическая библиотека, построенная по принципу управления объектами математического знания, а не математическими документами. В ее основу заложен фундаментальный принцип WDML — принцип создания сети математической информации, которая основана на знаниях, содержащихся в публикациях, представленных в электронных коллекциях. При этом электронная математическая библиотека рассматривается как большой открытый набор математической библиографической информации и математических понятий (аксиом, определений, теорем, доказательств, формул, уравнений, чисел, множеств, функций и др.), а также объектов (например, групп, колец), агрегированных из различных источников [8].

При проектировании цифровой библиотеки Lobachevskii-DML использовались полученные нами ранее результаты по управлению математическими знаниями [4–6], а также разработанные методы структурного и семантического анализа математических документов [11–14]. Учен опыт, приобретенный при проектировании платформы Science.Tatarstan [15], а также использованы подходы, апробированные при внедрении в практику работы научных журналов информационной системы Open Journal System (OJS) [16, 17].

Сервисы библиотеки Lobachevskii-DML

Начиная с 2003 года, для навигации в архиве статей электронного журнала Lobachevskii Journal of Mathematics (LJM) нами были разработаны сервисы, основанные на технологиях Семантического Веба [18, 19]; в соответствии со стандартом RDF построена иерархическая модель метаданных публикаций. В электронной коллекции LJM за 1998–2007 годы, размещенной в Lobachevskii-DML, реализован поиск по формулам, основанный на методе поиска по MathML-представлению документов, предложенному в [20–22]. Для расширения набора документов, обрабатываемых поисковым сервисом, ведутся исследования по автоматизации процессов семантического представления документов.

Предложена система сервисов, предоставляющих веб-инструменты подготовки публикации в TeX-формате [23]. Сервисы организованы в виде конструкторов, позволяющих в интерактивном режиме оформить статью в соответствии со стилиевыми требованиями журнала, выполнить проверку на соответствие стилиевым требованиям научного журнала, снабдить статью ключевыми словами и оформить библиографию по специальному шаблону. Сервис подготовки списка литературы позволяет оформить библиографию в соответствии с правилами научных журналов, представленных в электронной библиотеке. Имеется также возможность в интерактивном режиме выбрать тип публикации и затем ввести выходные данные в поля формы. Кроме того, предусмотрены возможности конвертации библиографических данных, оформленных в системе BibTeX, а также размеченных командами пакета amsbib (разработанного в рамках проекта MathNet), а также выбора общепринятого сокращения названия журнала из сформированного перечня сокращений большинства научных журналов.

Работа выполнена за счет средств субсидии, выделенной Казанскому федеральному университету для выполнения государственного задания в сфере научной деятельности, проект 1.2368.2017/ПЧ, и при частичной финансовой поддержке РФФИ и Правительства Республики Татарстан в рамках научных проектов №№ 15-07-08522, 15-47-02472.

Источники:

- [1] Lesk M. Understanding Digital Library. 2nd Ed. Elsevier Inc., 2005. 424 p.
- [2] Bouche T. Digital Mathematics Libraries: The good, the bad, the ugly // Mathematics in Computer Science. 2010. No 3. P. 227–241. DOI: 10.1007/s11786-010-0029-2.
- [3] Carette J., and Farmer W.M. A Review of Mathematical Knowledge Management. // Carette J. et al. (eds.) Intelligent Computer Mathematics. LNCS. 2009. V. 5625. P. 233–246. DOI: 10.1007/978-3-642-02614-0_21.
- [4] Elizarov A.M., Kirillovich A.V., Lipachev E.K., Nevzorova O.A., Solovyev V.D., and Zhiltsov N.G. Mathematical knowledge representation: semantic models and formalisms // Lobachevskii Journal of Mathematics. 2014. V. 35. No 4. P. 348–354. DOI: 10.1134/S1995080214040143.
- [5] Елизаров А.М., Кириллович А.В., Липачев Е.К., Невзорова О.А. Управление математическими знаниями: онтологические модели и цифровые технологии // Аналитика и управление данными в областях с интенсивным использованием данных. XVIII межд. конф. DAMDID/RCDL'2016. М.: ФИЦ ИУ РАН, 2016. С. 95–101 (CEUR Workshop Proceedings. 2016. V. 1752. P. 44–50. [Электр. ресурс]. URL: <http://ceur-ws.org/Vol-1752/paper08.pdf>).

- [6] 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 // Доклады Академии наук. 2016. Т. 467 (4). С. 392–395. DOI: 10.7868/S0869565216100042.
- [7] Sylwestrzak W., Borbinha J., Bouche T., Nowinski A., and Sojka P. EuDML – Towards the European Digital Mathematics Library. // Sojka, P. (ed.) Towards a Digital Mathematics Library. Paris: Masaryk University Press, 2010. P. 11–26. [Электр. ресурс]. URL: http://dml.cz/bitstream/handle/10338.dmlcz/702569/DML_003-2010-1_5.pdf.
- [8] Developing a 21st Century Global Library for Mathematics Research. Washington: The National Academies Press, 2014. 131 p. [Электр. ресурс]. DOI: <https://doi.org/10.17226/18619>.
- [9] Chebukov D.E., Izaak A.D., Misyurina O.G., Pupyrev Yu.A., and Zhizhchenko A.B. Math-Net.Ru as a Digital Archive of the Russian Mathematical Knowledge from the XIX Century to Today. Intelligent Computer Mathematics // LNCS. 2013. V. 7961. P. 344–348. DOI: 10.1007/978-3-642-39320-4_26.
- [10] Bartošek M., Lhoták M., Rákosník J., Sojka P., and Šárfy M. The DML-CZ Project: Objectives and First Steps. // Borwein J.M., Rocha E.M., Rodrigues J.F. (eds.) Communicating Mathematics in the Digital Era. A K Peters, Ltd., 2008. P. 75–86.
- [11] Елизаров, А.М., Зуев, Д.С., Липачев, Е.К., Малахальцев, М.А. Сервисы структурирования математического контента и интеграция электронных математических коллекций в научное информационное пространство // Электронные библиотеки: перспективные методы и технологии, электронные коллекции. Тр. XIV Всерос. науч. конф. RCDL-2012. Переславль-Залесский, 15–18 октября 2012 года. Переславль-Залесский, 2012. С. 309–312 (CEUR Workshop Proceedings. 2012. V. 934. P. 309–312. [Электр. ресурс]. URL: <http://ceur-ws.org/Vol-934/paper47.pdf>).
- [12] Елизаров А.М., Липачёв Е.К., Хохлов Ю.Е. Семантические методы структурирования математического контента, обеспечивающие расширенную поисковую функциональность // Информационное общество. 2013. № 1–2. С. 83–92.
- [13] Биряльцев Е.В., Елизаров А.М., Жильцов Н.Г., Липачёв Е.К., Невзорова О.А., Соловьев В.Д. Методы анализа семантических данных математических электронных коллекций // Научно-техническая информация. Сер. 2. Информационные процессы и системы. 2014. №4. С. 12–17.
- [14] Елизаров А.М., Липачёв Е.К., Невзорова О.А., Соловьев В.Д. Методы и средства семантического структурирования электронных математических документов // Доклады Академии наук. 2014. Т.457(6). С. 642–645. DOI: 10.7868/S0869565214240049.
- [15] Ахметов Д.Ю., Герасимов А.Н., Грачев А.О., Елизаров А.М., Липачёв Е.К. Облачная платформа поддержки электронных научных изданий // Ученые записки Института социальных и гуманитарных знаний. 2014. № 1-1 (12). С. 13–19.

- [16] Елизаров А.М., Зуев Д.С., Липачёв Е.К. Информационные системы управления электронными научными журналами // Научно-техническая информация. Сер. 1: Организация и методика информационной работы. 2014. № 3. С. 31–38. DOI: 10.3103/S0147688214010109.
- [17] Elizarov A., Lipachev E., Zuev D. Mathematical content semantic markup methods and open scientific e-journals management systems // Communications in Computer and Information Science. 2014. V. 468. P. 242–251. DOI: 10.1007/978-3-319-11716-4_22.
- [18] Глухов В.А., Елизаров А.М., Липачёв Е.К., Малахальцев М.А. Электронные научные издания: переход на технологии Семантического Веба // Электронные библиотеки. 2007. Т. 10. № 1.
- [19] Веселаго В.Г., Елизаров А.М., Липачёв Е.К., Малахальцев М.А. Формирование и поддержка физико-математических электронных научных изданий: переход на технологии Семантического Веба. // Научно-исследовательский институт математики и механики им. Н.Г. Чеботарева Казанского государственного университета. 2003–2007 гг. Казань: Изд-во Казан. ун-та, 2008. С. 456–476.
- [20] Елизаров А.М., Липачев Е.К., Малахальцев М.А. Веб-технологии для математика. Основы MathML. М.: Физматлит, 2010. 198 с.
- [21] Елизаров А.М., Липачёв Е.К., Малахальцев М.А. Технологии MathML поиска по формулам в электронных математических коллекциях // Межд. науч.-практ. конф. ИТОН-2012. Казань: Изд-во Казан. ун-та, 2012. С. 85–86.
- [22] Елизаров А.М., Липачев Е.К., Малахальцев М.А. Сервисы электронных естественнонаучных коллекций, построенные на основе технологии MathML // Научный сервис в сети Интернет: суперкомпьютерные центры и задачи. М.: МГУ, 2010. С. 533–534.
- [23] Ахметов Д.Ю., Елизаров А.М., Липачев Е.К. Информационные системы и сервисы комплексной поддержки периодических научных изданий // Научный сервис в сети Интернет. Тр. XVII Всерос. науч. конф. М.: ИПМ им. М.В. Келдыша, 2015. С. 16–25.