

А.М. Елизаров, Е.К. Липачёв

**Семантические методы и инструменты электронной математической библиотеки Lobachevskii-DML.**

***Рекомендуемая форма библиографической ссылки***

Елизаров А.М., Липачёв Е.К. Семантические методы и инструменты электронной математической библиотеки Lobachevskii-DML. // Научный сервис в сети Интернет: труды XIX Всероссийской научной конференции (18-23 сентября 2017 г., г. Новороссийск). — М.: ИПМ им. М.В.Келдыша, 2017. — С. 130-136. — URL: <http://keldysh.ru/abrau/2017/73.pdf>  
doi:[10.20948/abrau-2017-73](https://doi.org/10.20948/abrau-2017-73)

Размещена также [презентация к докладу](#)

# Семантические методы и инструменты электронной математической библиотеки Lobachevskii-DML

А.М. Елизаров, Е.К. Липачёв

*Институт математики и механики им. Н.И. Лобачевского  
Казанского (Приволжского) федерального университета*

**Аннотация.** Представлены сервисы управления научным контентом, разработанные при создании электронной математической библиотеки Lobachevskii-DML. Описаны технологии поддержки новых форм научных публикаций, взаимодействия с информационными системами научных журналов, электронными математическими архивами и библиотеками.

**Ключевые слова:** электронные документы, электронные библиотеки, электронные математические библиотеки, digital mathematical library, DML, WDML, электронная библиотека Lobachevskii DML.

**Введение.** Создание и развитие специализированных электронных библиотек (Digital Library – DL) являются одним из направлений формирования глобальной научной инфраструктуры [1]. В области математики проблемы интеграции знаний, полученных за весь «печатный» период развития этой науки, рассматривались в ряде проектов (например, [2–4]). Даже если проекты носили локальный характер, методы и инструменты, разрабатываемые в ходе их выполнения, были ориентированы на всеобъемлющую интеграцию знаний (см., например, [5]). Текущее видение задач по формированию глобальной инфраструктуры математических знаний оформлено в документах проекта Всемирной цифровой математической библиотеки – World Digital Mathematics Library (WDML) [6]. При этом отмечено, что ведущая роль в формировании электронных математических коллекций, создании методов проверки, сопровождении метаданными, аннотировании и др. отводится «меньшим» DML.

Настоящая работа посвящена представлению электронной библиотеки Lobachevskii Digital Mathematics Library (Lobachevskii-DML, <http://www.lobachevskii-dml.ru/>), которая разрабатывается нами в соответствии с основными принципами WDML и задачей которой является разработка инструментов управления математическим контентом, учитывающих не только специфику математических текстов, но особенности обработки русскоязычных текстов. Частной задачей при построении этой электронной библиотеки является интеграция математических ресурсов Казанского университета, что объясняет выбор ее названия. Кроме того, проект создания Lobachevskii-DML стартовал в год 225-летия

со дня рождения гениального математика Николая Ивановича Лобачевского, основателя неевклидовой геометрии, воспитанника и ректора Казанского университета (с 1827 по 1845 годы). 2017 год объявлен в Казанском университете «Годом Н.И. Лобачевского».

## **1. Объекты и связи между ними – основа управления электронным контентом**

Управление электронными математическими документами – уникальная и сложная задача. Оно связано как с обработкой математических формул, так и со специфической структурой математического документа, состоящего из логически связанной последовательности определений, теорем, доказательств и ссылок. Ключевой идеей, обозначенной в документах проекта WDML, является разработка классов объектов для адекватного описания и исследования математического содержания: предложена новая парадигма представления электронного математического контента, основой которой являются элементы (классы) и их взаимосвязи. Выделение классов математических объектов и формирование на их основе онтологий областей знаний позволят создать новые инструменты для обработки информации, в частности, извлечения и обработки формул, поиску похожих результатов и т. п.

В наших работах [7–9] предложены методы структурного анализа математических документов и выделения объектов из них. В [10, 11] описана цифровая экосистема OntoMath, состоящая из онтологий, инструментов текстовой аналитики и приложений, предназначенная для управления математическими знаниями. Семантическое аннотирование математических текстов базируется на онтологии, построенной в рамках проекта Mocassin, и онтологии профессиональной математики OntoMathPRO [12]. Важным приложением, разработанным на основе этих онтологий, является специальная программная платформа для подготовки математического набора связанных данных для публикации в облаке LOD. Другим важным инструментом служит семантический сервис поиска по математическим формулам [11]. Еще одно приложение экосистемы OntoMath – рекомендательная система для коллекций физико-математических документов, позволяющая, в частности, построить список «близких» документов для заданного документа на основе объектов (см. [13]). Указанные инструменты включены в состав сервисов электронной библиотеки Lobachevskii-DML.

## **2. Поддержка новых форм публикаций в электронных библиотеках**

В настоящее время активно формируется новая среда научно-образовательной деятельности, основанная на использовании интернет-технологий. В связи с этим, наряду с традиционными формами научного обмена, ориентированными, в основном, на печатные издания или их электронные версии, появляются новые формы научных коммуникаций. Как наиболее заметные, отметим цифровые презентации, ставшие необходимым атрибутом докладов на научных

конференциях, научные форумы и блоги, электронные препринты, вебинары и видеолекции [14]. Интернет-активность ученого, часто рассматриваемая как его обязанность (см., например, [15]), предполагает привлечение всех возможных средств коммуникаций. Использование новых форм научного обмена не должно нарушать сложившихся традиций научного сообщества, предусматривающих оценку научной работы в виде рецензирования, систему цитирования и т. д. Как следствие, научный документ любой формы должен иметь библиографическое описание и набор метаданных. Например, для «живых публикаций» предложено включать в описание дату последней редакции [16]. Удачным примером реализации модели «живых публикаций» является Стэнфордская философская энциклопедия (<https://plato.stanford.edu/>). Все статьи этой энциклопедии написаны специалистами в соответствующих дисциплинах и прошли процедуру рецензирования. Авторы поддерживают статьи в актуальном состоянии, периодически обновляя их. Каждое обновление статьи снова подвергается процедуре рецензирования, причем ведется история версий публикации (имеется возможность сослаться как на последнюю, так и на любую из предыдущих версий статьи).

Другим средством поддержки интернет-активности ученого являются блоги (см., например, [17, 18]). В качестве примера отметим блог современного математика Стивена Вольфрама (<http://blog.stephenwolfram.com/>). Другой пример – блог проекта WDML (<https://blog.wias-berlin.de/imu-icm-panel-wdml/>). Блоги можно использовать как средство организации открытого научного рецензирования. Такое рецензирование, в дополнение к традиционному, позволяет избежать конфликта интересов и привлечь к экспертизе более широкий круг экспертов. Одной из разновидностей открытого рецензирования является краудсорсинг-рецензирование, при котором в процессе рецензирования может принять участие любой представитель научного сообщества. Краудсорсинг-рецензирование можно организовать с помощью блога, как это было сделано, например, при обсуждении статьи о предполагаемом решении классической проблемы о равенстве классов P и NP (<https://rjlipton.wordpress.com/2010/08/12/fatal-flaws-in-deolalikars-proof/>; <https://rjlipton.wordpress.com/2010/08/09/issues-in-the-proof-that-p-neq-np/>).

### **3. Электронная математическая библиотека Lobachevskii-DML**

Lobachevskii-DML (<http://www.lobachevskii-dml.ru/>) – электронная математическая библиотека, построенная в соответствии с основными рекомендациями, указанными в документах WDML [6, 19] (см. рис. 1).

В настоящее время контент Lobachevskii-DML образован из следующих электронных коллекций:

- статей журнала Lobachevskii Journal of Mathematics (LJM) за 1998–2007 годы с семантическими инструментами навигации, включая поиск по формулам (см. рис. 2);
- статей журнала LJM с 2007 года по настоящее время, связанных с электронными ресурсами издательства Springer;

- «Трудов Математического центра имени Н.И. Лобачевского»;
- отдельных изданий трудов конференций по математике и механике, проведенных в Казанском университете с 1994 года по настоящее время;
- статей журнала «Известия высших учебных заведений. Математика»;
- статей журнала «Учёные записки Казанского университета. Серия Физико-математические науки».

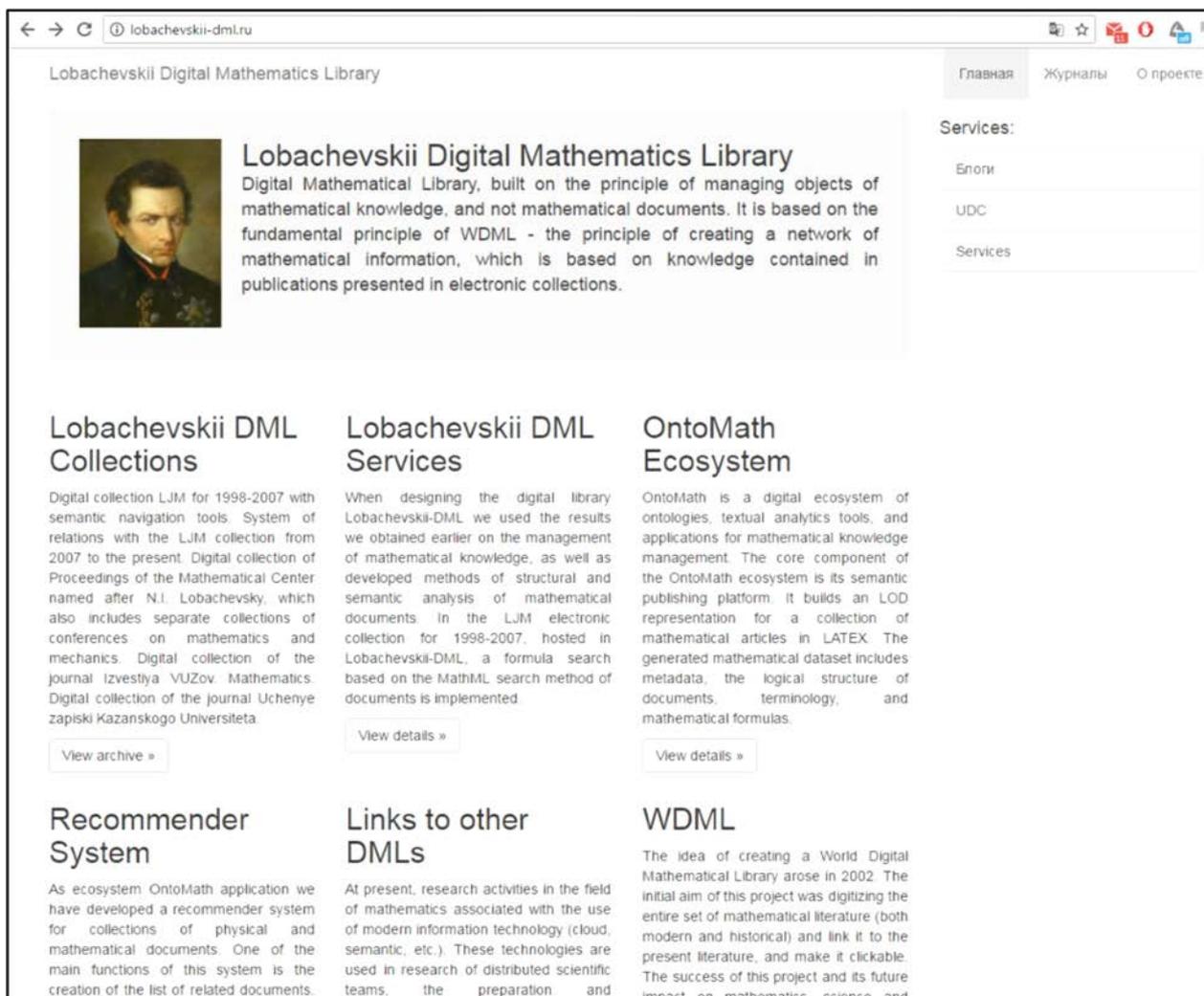


Рис. 1. Главная страница цифровой библиотеки Lobachevskii-DML

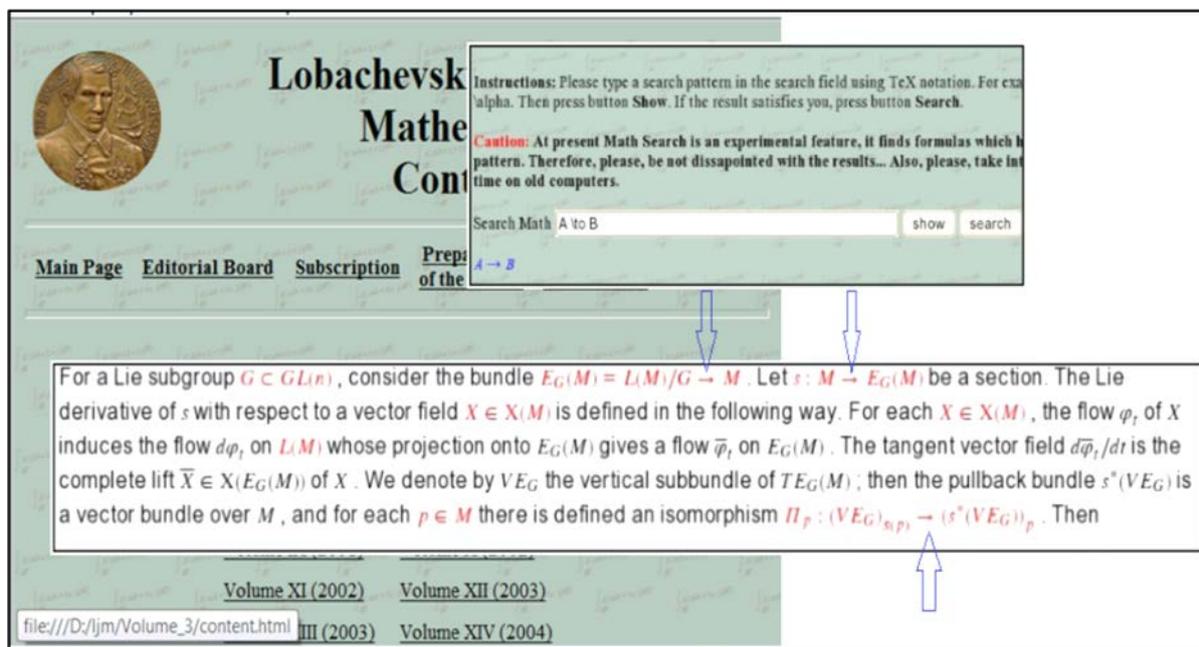


Рис. 2. Поиск по формульным фрагментам

При проектировании цифровой библиотеки Lobachevskii-DML использовались результаты по управлению математическими знаниями, полученные нами ранее, а также разработанные методы структурного и семантического анализа математических документов [7–13].

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 Российского фонда фундаментальных исследований и Правительства Республики Татарстан в рамках научных проектов №№ 15-07-08522, 15-47-02472.

### Литература

1. Lesk M. Understanding Digital Library. Second Edition. — Elsevier Inc., 2005. — 424 p.
2. Sylwestrzak W., Borbinha J., Bouche T., Nowinski A., and Sojka P. EuDML – Towards the European Digital Mathematics Library. In: Sojka, P. (ed.) Towards a Digital Mathematics Library. – Paris: Masaryk University Press, 2010. — P. 11-26. [http://dml.cz/bitstream/handle/10338.dmlcz/702569/DML\\_003-2010-1\\_5.pdf](http://dml.cz/bitstream/handle/10338.dmlcz/702569/DML_003-2010-1_5.pdf).
3. 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.
4. 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 //

- In: Carette J., Aspinall D., Lange C., Sojka P., Windsteiger W. (eds) Intelligent Computer Mathematics. CICM 2013. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg. — 2013. — V. 7961. — P. 344-348. doi: 10.1007/978-3-642-39320-4\_26.
5. Bartošek M., Lhoták M., Rákosník J., Sojka P., and Šárky M. The DML-CZ Project: Objectives and First Steps. In Borwein J.M., Rocha E.M., Rodrigues J.F. (eds.) Communicating Mathematics in the Digital Era. — A K Peters, Ltd., 2008. — P. 75-86.
  6. Developing a 21st Century Global Library for Mathematics Research. Washington, The National Academies Press, 2014. — 131 p. doi: 10.17226/18619.
  7. 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.
  8. Биряльцев Е.В., Елизаров А.М., Жильцов Н.Г., Липачёв Е.К., Невзорова О.А., Соловьёв В.Д. Методы анализа семантических данных математических электронных коллекций // Научно-техническая информация. Сер. 2. Информационные процессы и системы. — 2014. — № 4. — С. 12-17.
  9. Елизаров А.М., Липачёв Е.К., Невзорова О.А., Соловьёв В.Д. Методы и средства семантического структурирования электронных математических документов // Доклады Академии наук. — 2014. — Т. 457 (6). — С. 642-645. doi: 10.7868/S0869565214240049.
  10. 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 международная конференция. — 2016. — С. 95-101. (CEUR Workshop Proceedings. — 2016. — V. 1752. — P. 44-50. <http://ceur-ws.org/Vol-1752/>).
  11. Elizarov A., Kirillovich A., Lipachev E., and Nevzorova O. Digital Ecosystem OntoMath: Mathematical Knowledge Analytics and Management // Communications in Computer and Information Science, Springer. — 2017. — V. 706. — P. 33-46. doi: 10.1007/978-3-319-57135-5\_3.
  12. Елизаров А.М., Жильцов Н.Г., Кириллович А.В., Липачев Е.К. Семантическое аннотирование в системе управления физико-математическим контентом // Научный сервис в сети Интернет: труды XVII Всероссийской научной конференции. ИПМ им. М.В. Келдыша. — 2015. — С. 98-103. <http://keldysh.ru/abrau/2015/119.pdf>.
  13. 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.

14. Изаак А.Д., Мисюрина О.Г., Пупырев Ю.А., Чебуков Д.Е. Индексация видеозаписей научных мероприятий в Общероссийском математическом портале Math-Net.Ru // Научный сервис в сети Интернет: труды XVII Всероссийской научной конференции. ИПМ им. М.В. Келдыша. — 2015. — С. 114-118. <http://keldysh.ru/abrau/2015/proc.pdf#page=114>
15. Горбунов-Посадов М.М. Интернет-активность как обязанность ученого. Редакция от 25.02.2017. <http://keldysh.ru/gorbunov/duty.htm>.
16. Горбунов-Посадов М.М., Скорнякова Р.Ю. Дата последней редакции как живой атрибут живой публикации // Научный сервис в сети Интернет: труды XVIII Всероссийской научной конференции (19–24 сентября 2016 г., г. Новороссийск). — М.: ИПМ им. М.В. Келдыша, 2016. — С. 113-114. — doi: 10.20948/abrau-2016-48. <http://keldysh.ru/abrau/2016/48.pdf>
17. Puschmann C. (Micro)Blogging Science? Notes on Potentials and Constraints of New Forms of Scholarly Communication // In Bartling S. and Friesike S. (eds.). *Opening Science*. — Springer Cham, 2014. — P. 89-106. doi: 10.1007/978-3-319-00026-8\_6.
18. Елизаров А.М., Кириллович А.В., Липачёв Е.К. Блоги в системе научных коммуникаций // Ученые записки Института социальных и гуманитарных знаний. — 2017. — № 1 (15). — С. 209-214.
19. Елизаров А.М., Кириллович А.В., Липачёв Е.К. Структура и сервисы цифровой математической библиотеки Lobachevskii-DML // Ученые записки Института социальных и гуманитарных знаний. — 2017. — № 1 (15). — С. 215-220.