

ЦИФРОВЫЕ МАТЕМАТИЧЕСКИЕ БИБЛИОТЕКИ В ОБРАЗОВАТЕЛЬНОМ ИНФОРМАЦИОННОМ ПРОСТРАНСТВЕ

Елизаров А. М.¹ [0000-0003-2546-6897], Липачёв Е. К.² [0000-0001-7789-2332]

*Институт информационных технологий и интеллектуальных систем
Казанского федерального университета, ул. Кремлевская, 35, г. Казань, 420008*

¹amelizarov@gmail.com, ²elipachev@gmail.com

Аннотация

Предложены методы управления электронными математическими коллекциями и их использования в образовательном процессе. Электронные коллекции формируются с помощью программных инструментов цифровой математической библиотеки Lobachevskii-DML, которая рассматривается как составной элемент единого цифрового пространства математических знаний. Логическая структура, присущая математическим документам, используется в методах управления контентом электронных коллекций.

Ключевые слова: *электронные математические коллекции, цифровые математические библиотеки, цифровая математическая библиотека Lobachevskii-DML, единое цифровое пространство математических знаний.*

Использование знаний, извлеченных в результате поиска в Сети, является повседневной практикой современных научных исследований и образовательной деятельности. В связи с этим возникает ряд проблем, наиболее важной из которых является верификация информации, полученной из интернет-источников, – в хорошо известном стеке стандартов Семантического веба уровень “Trust” («Доверительная информация») размещен на вершине стека и выделен как наиболее сложный (см., например, [1]). В определенной степени проблема доверия к образовательной и научной информации из Сети решается в формате единого пространства научных знаний (см., например, [2]).

В проектах формирования единого цифрового пространства математических знаний определяющая роль отведена цифровым математическим библиотекам. Одна из основных задач последних состоит в создании методов включения в

единое цифровое научное пространство как современных математических знаний, так и математических знаний, содержащихся в статьях и книгах, опубликованных в доцифровой период (см., например, [3]).

Цифровые математические библиотеки с помощью сервисов своих информационных платформ предоставляют доступ к электронным коллекциям научных и образовательных документов. Проекты построения цифровых математических библиотек не ограничиваются созданием электронных коллекций и также предполагают разработку методов извлечения и нормализации метаданных, необходимых для интеграции в единое научное пространство. Важной составляющей многих проектов являются реализация и предоставление на своих информационных платформах интеллектуальных программных инструментов, обеспечивающих работу с контентом, в частности, поисковые сервисы, учитывающие семантику предметной области (см., например, [4, 5]).

В рамках проекта по созданию цифровой математической библиотеки Lobachevskii-DML (<https://lobachevskii-dml.ru/>) с 2017 года сформированы электронные коллекции, включающие как статьи современного периода, так и ретро-документы [6]. Одним из этапов создания коллекций является формирование наборов метаданных, обеспечивающих функционирование сервисов цифровой библиотеки. Этот этап включает также создание методов нормализации метаданных по xml-схемам агрегирующих библиотек. Специфика математического контента потребовала корректировки известных и создания новых методов обработки документов – разработаны методы автоматической обработки больших массивов документов и формирования из них электронных коллекций с обязательным набором метаданных. Методы извлечения метаданных из документов коллекции основаны на анализе структуры и стилистических особенностей документов. Для уточнения и пополнения метаданных документов создана система извлечения знаний из Wikidata. При извлечении терминов из документов использованы известные подходы разрешения лексической многозначности именованных сущностей. Отметим также методы поиска близких документов, дублирующих документов (включая переведенных на другие языки), а также методы тематической кластеризации документов. Реализованные программные инструменты объединены в фабрику метаданных, обеспечивающую автоматизацию большинства рабочих процессов. Развитие методов работы с математическим контентом, вклю-

чая перечисленные выше, реализовано в образовательном процессе в рамках выполнения магистерских диссертаций. Часть разработанных методов подтверждена свидетельствами о регистрации программ для ЭВМ и баз данных и предоставлена в виде сервисов цифровой математической библиотеки Lobachevskii-DML. В частности, это «Программный комплекс выделения метаданных из коллекций физико-математических документов, представленных в формате OpenXML», «Программный комплекс автоматизированного формирования оригинал-макетов научных изданий из документов в офисных форматах», «Программный комплекс формирования рекомендаций по подбору рецензентов для научных документов в информационных издательских системах», «Программа автоматизированного формирования метаданных в формате Российского индекса научного цитирования», «Программный комплекс автоматического пре-процессинга цифровых коллекций математических документов» и «Программа нормализации метаданных в форматах инфометрических баз данных».

Отдельного внимания заслуживает цифровая экосистема OntoMath, представленная на информационной платформе цифровой библиотеки Lobachevskii-DML. Сервисы этой экосистемы формируют новые образовательные возможности, влияющие на эффективность процесса обучения математике [7, 8].

С точки зрения использования в образовательном процессе выделим следующие компоненты этой цифровой экосистемы: образовательная математическая онтология OntoMath^{Edu}, онтология профессиональной математики OntoMath^{PRO}, рекомендательная система поиска близких документов по математике, основанная на онтологиях математического знания, сервис терминологического аннотирования, поисковые сервисы по контенту электронных библиотек, учитывающие специфику математических документов, а также семантические сервисы поиска по математическим формулам.

Благодарности

Работа выполнена при финансовой поддержке Российского научного фонда (проект № 21-11-00105).

СПИСОК ЛИТЕРАТУРЫ

1. *Berners-Lee T.* Semantic Web – XML2000. Talk. December 2000. URL: <http://www.w3.org/2000/Talks/1206-xml2k-tbl/>.
2. *Каленов Н.Е., Савин Г.И., Сотников А.Н.* Архитектура единого цифрового

пространства научных знаний // Единое цифровое пространство научных знаний: проблемы и решения: сборник научных трудов / под ред. Н.Е. Каленова, А.Н. Сотникова. Москва; Берлин: Директмедиа Паблишинг, 2021. С. 7–16.

3. Developing a 21st Century Global Library for Mathematics Research, DC: The National Academies Press, Washington. doi:10.17226/18619.
4. *Elizarov A.M., Lipachev E.K., Zuev D.S.* Digital Mathematical Libraries: Overview of Implementations and Content Management Services // CEUR Workshop Proceedings. 2017. V. 2022. P. 317–325.
5. *Borwein J.M., Rocha E.M., Rodrigues J.F. (Eds.)* Communicating Mathematics in the Digital Era. A K Peters, Ltd., CRC Press, 2008.
6. *Elizarov A.M., Lipachev E.K.* Big Math Methods in Lobachevskii-DML Digital Library // CEUR Workshop Proceedings. 2019. Vol. 2523. P. 59–72.
7. *Elizarov A., Kirillovich A., Lipachev E., Nevzorova O.* Digital Ecosystem OntoMath: Mathematical Knowledge Analytics and Management // Communications in Computer and Information Science, Springer. 2017. Vol. 706. P. 33–46. doi:10.1007/978-3-319-57135-5_3.
8. *Kirillovich A., Shakirova L., Falileeva M., Lipachev E.* Towards an Educational Mathematical Ontology. L. Gómez Chova, et al. (Eds). Proc. of the 13th International Technology, Education and Development Conference (INTED2019), Valencia, Spain, March 11–13, 2019. IATED, P. 6823–6829 (2019).

DIGITAL MATHEMATICAL LIBRARIES IN EDUCATIONAL INFORMATION SPACE

Alexander Elizarov¹ [0000-0003-2546-6897], Evgeny Lipachev² [0000-0001-7789-2332]

Institute of Information Technologies and Intelligent Systems, Kazan (Volga Region) Federal University, ul. Kremlyovskaya, 35, Kazan, 420008

¹amelizarov@gmail.com, ²elipachev@gmail.com

Abstract

Methods for managing electronic mathematical collections and their use in the educational process are proposed. Electronic collections are formed using the software tools of the Lobachevskii-DML digital mathematical library, which is considered as an integral element of the common digital space of mathematical knowledge. The logical structure inherent in mathematical documents is used in methods for managing the content of electronic collections.

Keywords: *electronic mathematical collections, digital mathematical libraries, digital mathematical library Lobachevskii-DML, common digital space of mathematical knowledge.*

REFERENCES

1. *Berners-Lee T.* Semantic Web – XML2000. Talk. December 2000. URL: <http://www.w3.org/2000/Talks/1206-xml2k-tbl/>.
2. *Kalenov N.E., Savin G.I., Sotnikov A.N.* The Architecture of the Common Digital Space of Scientific Knowledge. In: N.E. Kalenov, A.N. Sotnikov (Eds.). *Edinoe cifrovoe prostranstvo nauchnyh znaniy: problemy i resheniya: sbornik nauchnyh trudov.* Direktmedia Publishing. Moscow, Berlin, 2021. P. 7–16. <https://doi.org/10.23681/610687>.
3. *Developing a 21st Century Global Library for Mathematics Research*, DC: The National Academies Press, Washington. doi:10.17226/18619.
4. *Elizarov A.M., Lipachev E.K., Zuev D.S.* Digital Mathematical Libraries: Overview of Implementations and Content Management Services // CEUR Workshop Proceedings. 2017. V. 2022. P. 317–325.
5. *Borwein J.M., Rocha E.M., Rodrigues J.F. (Ed.)* Communicating Mathematics in the Digital Era. A K Peters, Ltd., CRC Press, 2008.
6. *Elizarov A.M., Lipachev E.K.* Big Math Methods in Lobachevskii-DML Digital

Library // CEUR Workshop Proceedings. 2019. Vol. 2523. P. 59–72.

7. *Elizarov A., Kirillovich A., Lipachev E., Nevzorova O.* Digital Ecosystem OntoMath: Mathematical Knowledge Analytics and Management // Communications in Computer and Information Science, Springer. 2017. Vol. 706. P. 33–46. doi:10.1007/978-3-319-57135-5_3.
8. *Kirillovich A., Shakirova L., Falileeva M., Lipachev E.* Towards an Educational Mathematical Ontology. L. Gómez Chova, et al. (eds). Proc. of the 13th International Technology, Education and Development Conference (INTED2019), Valencia, Spain, March 11–13, 2019. IATED, 6823–6829 (2019).

СВЕДЕНИЯ ОБ АВТОРАХ



ЕЛИЗАРОВ Александр Михайлович – доктор физико-математических наук, профессор, заслуженный деятель науки Республики Татарстан, профессор кафедры программной инженерии Института информационных технологий и интеллектуальных систем Казанского (Приволжского) федерального университета.

Alexander Michailovich ELIZAROV – Doctor of Physics and Mathematics, Professor, Honoured Worker of Science of the Republic of Tatarstan, Kazan Federal University.

Current scientific interests: data mining, recommender systems, cloud computing, knowledge extraction technologies.

email: amelizarov@gmail.com;

ORCID: 0000-0003-2546-6897



ЛИПАЧЁВ Евгений Константинович – кандидат физико-математических наук, доцент, доцент кафедры Интеллектуальных технологий поиска Института информационных технологий и интеллектуальных систем Казанского (Приволжского) федерального университета.

Evgeny Konstantinovich LIPACHEV – Candidate of Physics and Mathematics, Associate Professor, Kazan Federal University.

Current scientific interests: Natural Language Processing, recommender systems, ontological modelling, knowledge extraction technologies.

URL: <https://kpfu.ru/Evgeny.Lipachev>.

email: elipachev@gmail.com;

ORCID: 0000-0001-7789-2332

Материал поступил в редакцию 23 февраля 2022 года