



П.О. Гафурова, А.М. Елизаров,
Е.К. Липачев, Д.М. Хамматова

**Методы формирования и
нормализации метаданных в
цифровой математической
библиотеке**

Рекомендуемая форма библиографической ссылки

Гафурова П.О., Елизаров А.М., Липачев Е.К., Хамматова Д.М. Методы формирования и нормализации метаданных в цифровой математической библиотеке // Научный сервис в сети Интернет: труды XXI Всероссийской научной конференции (23-28 сентября 2019 г., г. Новороссийск). — М.: ИПМ им. М.В.Келдыша, 2019. — С. 234-244. — URL: <http://keldysh.ru/abrau/2019/theses/28.pdf> doi:[10.20948/abrau-2019-28](https://doi.org/10.20948/abrau-2019-28)

Размещена также [презентация к докладу](#)

Методы формирования и нормализации метаданных в цифровой математической библиотеке

П.О. Гафурова¹, А.М. Елизаров^{1,2}, Е.К. Липачев^{1,2}, Д.М. Хамматова¹

¹*Институт математики и механики им. Н.И. Лобачевского*

²*Высшая школа информационных технологий и интеллектуальных систем
Казанского (Приволжского) федерального университета*

Аннотация. Предложены и реализованы методы интеграции электронных математических коллекций Казанского университета в отечественные и зарубежные цифровые математические библиотеки. Представлены алгоритмы пополнения электронных коллекций цифровой математической библиотеки Lobachevskii-DML и формирования метаданных документов коллекций в соответствии с синтаксическими правилами этой цифровой библиотеки. Разработаны сервисы нормализации метаданных коллекций цифровой библиотеки Lobachevskii-DML в соответствии с DTD-правилами и XML-схемами Journal Archiving and Interchange Tag Suite (NISO JATS) V1.0, V1.1, V1.2. Приведен алгоритм автоматизированной подготовки метаданных электронных коллекций библиотеки Lobachevskii-DML по правилам библиографической базы по компьютерным наукам “Dblp Computer Science Bibliography” (DBLP).

Ключевые слова: электронные математические коллекции, цифровые библиотеки, цифровые математические библиотеки, формирование и извлечение метаданных, семантические связи информационных объектов, сервисы нормализации метаданных, цифровая математическая библиотека Lobachevskii-DML.

Methods of Formation and Normalization of Metadata in the Digital Mathematical Library

P.O. Gafurova¹, A.M. Elizarov^{1,2}, E.K. Lipachev^{1,2}, D.M. Khammatova¹

¹*N. I. Lobachevskii Institute of Mathematics and Mechanics,*

²*Higher School of Information Technologies and Intelligent Systems,
Kazan (Volga Region) Federal University*

Abstract. We offer methods for integrating electronic mathematical collections of the Kazan University into domestic and foreign digital mathematical libraries. We present the algorithms for enriching the electronic collections of the Lobachevskii

Digital Mathematical Library (Lobachevskii-DML) and generating metadata for collection documents in accordance with the syntax rules of this digital library. We developed the services for normalization of the collection metadata of the Lobachevskii-DML in accordance with DTD rules and Journal Archiving and Interchange Tag Suite XML schemas (NISO JATS) V1.0, V1.1, V1.2. We present an algorithm for automated preparation of electronic collections of the Lobachevskii-DML library according to the rules of the bibliographic database on computer science “Dblp Computer Science Bibliography” (DBLP).

Keywords: electronic mathematical collections, digital libraries, digital mathematical library, formation and extraction of metadata, semantic links of information objects, metadata normalization services, Lobachevskii-DML.

1. Введение

С развитием информационно-коммуникационных технологий впервые появилась возможность сделать доступными научные знания, созданные за весь печатный период. Поэтому неслучайным является появление таких инициатив, как “World Digital Mathematics Library” (WDML, <https://www.mathunion.org/ceic/library/world-digital-mathematics-library-wdml>, – проект “Всемирной цифровой математической библиотеки”) и “The Global Digital Mathematics Library” (GDML – проект “Глобальной цифровой математической библиотеки”), нацеленных на разработку основных принципов интеграции научных знаний в области математики [1, 2]. Целью проекта “The European Digital Mathematics Library” (EuDML, <https://initiative.eudml.org/>, “Европейский проект интеграции математических ресурсов”) является интеграция математических ресурсов европейских цифровых библиотек [3, 4]. Российский проект MathNet.Ru (<http://www.mathnet.ru/>) сделал доступными архивы отечественных журналов и сборников, предложил методы навигации и расширенного поиска по математическому контенту, а также систему связей с библиографическими базами данных [5].

С точки зрения программных документов WDML (см. [1]) основополагающая роль в интеграции математических знаний отводится цифровым математическим библиотекам (см., напр., [6]). В рамках этих библиотек разрабатываются методы обработки и управления математическими документами, основанные на семантических связях не только между документами, но и с объектами, содержащимися в них [5, 7–11]. Одновременно в связи со значительным ростом объемов научных публикаций появилась необходимость создания специализированных методов автоматизированной обработки больших массивов документов [11–13].

К числу названных библиотек относится Lobachevskii Digital Mathematical Library (Lobachevskii-DML, <https://lobachevskii-dml.ru/>), создаваемая в Казанском университете [9, 10]. Следуя принципам WDML, при разработке этой библиотеки акцент сделан на использовании семантических методов обработки математического контента. Как один из результатов в этом направлении можно

рассматривать спроектированную и разрабатываемую в настоящее время цифровую экосистему OntoMath [8, 14].

В настоящей статье представлены методы формирования метаданных электронных коллекций библиотеки Lobachevskii-DML. Для решения задачи интеграции создаваемых электронных коллекций в агрегирующие цифровые библиотеки, такие, как EuDML, MathNet.Ru, DBLP, предложены методы преобразования метаданных по схемам, принятым в указанных библиотеках. Описаны методы нормализации метаданных коллекций цифровой библиотеки Lobachevskii-DML в соответствии с DTD-правилами и XML-схемами Journal Archiving and Interchange Tag Suite (NISO JATS, <https://jats.nlm.nih.gov/archiving/>) V1.0, V1.1, V1.2 [15]. Для обозначения методов формирования и преобразования метаданных документов в соответствии с правилами и XML-схемами цифровых библиотек и наукометрических баз данных мы используем термин *нормализация* (см. также [11]). Метод нормализации метаданных по правилам NISO JATS послужил основой формирования обязательного и фундаментального наборов метаданных EuDML. Также приведен алгоритм автоматизированной подготовки метаданных электронных коллекций библиотеки Lobachevskii-DML по правилам библиографической базы по компьютерным наукам “Dblp Computer Science Bibliography” (DBLP, <https://dblp.uni-trier.de/>).

2. Форматы метаданных документов цифровых математических библиотек

В настоящее время многие наукометрические базы данных осуществляют индексацию статей, опубликованных в ведущих математических журналах. Эти базы предъявляют различные требования к набору метаданных этих документов, а также схемам их представления (см., например, [16]). Отметим, что, как правило, не индексируются новые формы публикаций: презентации, научные блоги и видеолекции – важные компоненты современных цифровых библиотек.

Цифровые математические библиотеки при формировании коллекций, входящих в них, используют различные форматы метаданных. Это объясняется тем, что многие такие коллекции образуются из статей, опубликованных в академических журналах и оформленных в соответствии с правилами, установленными в этих журналах и отличающимися требованиями к используемым метаданным. Эти отличия касаются, прежде всего, состава метаданных и их формата и более всего заметны в архивных коллекциях научных журналов. При этом по постоянно расширяемому набору метаданных в них (даже за относительно короткий временной период) можно судить о возрастающей роли метаданных в совершенствовании научных коммуникаций в научно-исследовательской деятельности.

В качестве примера укажем архив статей журнала «Известия высших учебных заведений. Математика» («Russian Mathematics (Izvestiya VUZ. Matematika)»), <https://kpfu.ru/science/nauchnye-izdaniya/ivrm>, статьи которого со-

ставляют одну из коллекций цифровой библиотеки Lobachevskii DML. Названный журнал издается с 1957 года, а относительно полным набором метаданных сопровождаются только статьи, опубликованные в этом журнале, начиная с 2015 года. В статьях, опубликованных до 2008 года, отсутствуют ключевые слова и аннотации. Идентификатор ORCID автора статьи пока не указывается, а аффилиация авторов появилась только в статьях последних лет.

Как известно, о понятии «аффилиация», указывающей на место выполнения работы, результаты которой изложены в научной статье, российские авторы узнали, когда начали публиковаться в зарубежных изданиях. При этом аффилиация авторов научных публикаций используется в международных наукометрических базах данных как важный фактор для определения рейтинга журнала и показатель публикации активности научных учреждений, в которых ученые проводят свои исследования. ORCID (Open Researcher and Contributor ID) – это уникальный идентификационный код, позволяющий ликвидировать разночтения в фамилиях и именах авторов научных статей.

История расширения набора используемых метаданных, описанная в приведенном выше примере, является типичной практически для всех научных журналов. Поэтому возникает необходимость в разработке методов извлечения метаданных из документов, а также методов приведения уже созданных метаданных в форматы соответствующих наукометрических баз данных. Отметим также, что участие в таких проектах, как EuDML, предполагает предоставление наборов метаданных, сформированных по схемам агрегаторов математических ресурсов.

Одним из этапов интеграции электронных математических коллекций в EuDML является нормализация метаданных этих коллекций по правилам образования обязательного (obligatory) набора метаданных. Для описания статей из математических журналов EuDML применяет XML-схемы NISO JATS V1.0, а общая схема метаданных этой цифровой библиотеки описана в [17]. Метаданные разделены на обязательные (obligatory metadata), фундаментальные (fundamental metadata) и дополнительные (supplemental metadata) [18]. Минимальным из них по составу является обязательный набор метаданных, в который включены: название статьи на языке оригинала, список авторов, библиография, уникальный идентификатор статьи, например, doi и URL полного текста статьи. Фундаментальный набор метаданных дополнительно к обязательным метаданным включает аннотацию статьи и ключевые слова.

Отметим, что ряд электронных коллекций библиотеки Lobachevskii-DML физически размещен в других цифровых библиотеках. Например, журнальная коллекция «Известия вузов. Математика» оцифрована, снабжена метаописаниями и представлена на портале MathNet.Ru (<http://www.mathnet.ru/php/journal.phtml?jrnid=ivm>). Нашими задачами являются пополнение таких коллекций дополнительными метаданными, а также автоматическое выделение объектов и установление семантических связей между ними.

При формировании фундаментального набора метаданных электронных коллекций, хранящихся на внешних ресурсах, первоначально производится импорт метаданных, представленных на этих ресурсах. Для этой цели на языке C# и с использованием функций пакета HtmlAgilityPack (<https://html-agility-pack.net/>) разработана программа выделения метаданных с веб-страниц и их записи в XML-формате цифровой библиотеки Lobachevskii-DML, пополнения и последующего преобразования по схемам EuDML. Например, для коллекции «Известия вузов. Математика» выполнены следующие шаги.

Часть метаданных импортирована из блока «Цитирование в формате AMSBIB», представленного на веб-странице статьи портала MathNet.Ru. Затем с веб-страницы считаны ключевые слова и гиперссылка на страницу портала Springer Link (<https://link.springer.com/journal/11982>) с англоязычной версией статьи. Эта информация включена в состав метаданных, и выполнен переход по гиперссылке.

Следующий шаг включал анализ веб-страницы англоязычной версии статьи, извлечение и запись метаданных. Далее был сгенерирован персональный идентификатор этой статьи, который предложено создавать как объединение строк – идентификатора журнала (значение атрибута «jrnid=») и идентификатора статьи (значение атрибута «paperid=») на портале MathNet.Ru.

В настоящее время схемы, предложенные EuDML, не позволяют соединить в рамках одного метаописания статью, опубликованную на русском языке, и ее переводную версию на английском языке. Например, статья «Авхадиев Ф.Г., Насибуллин Р.Г., Шафигуллин И.К. // Известия высших учебных заведений. Математика. 2018. № 8. С. 88–92» и ее перевод на английский язык “Avkhadiev F.G., Nasibullin R.G., Shafigullin I.K. // Russian Mathematics. 2018. V. 62. No 8. P. 76–79” в фундаментальном наборе EuDML приходится описывать как различные статьи в разных журналах. В коллекциях Lobachevskii-DML, а также библиотеках eLibrary.ru и MathNet.ru такие статьи представлены как дубликаты одного документа.

Отметим, что процесс подготовки метаданных в формате eLibrary.ru автоматизирован (см. [19, 20]) и успешно используется в журнале «Электронные библиотеки».

3. Нормализация метаданных по схемам DBLP

Одной из самых популярных и авторитетных библиотек по компьютерным наукам является «Dblp Computer Science Bibliography» (DBLP, <https://dblp.uni-trier.de/>) [16]. Необходимым условием включения электронных коллекций в эту библиотеку является реорганизация и нормализация метаданных документов цифровой библиотеки. Среди коллекций цифровой библиотеки Lobachevskii DML требованиям DBLP по контенту удовлетворяет коллекция журнала «Электронные библиотеки» (“Russian Digital Libraries Journal”, <https://elbib.ru/>).

С 2015 года в этом журнале была использована новая модель представления документов и внедрена издательская система Open Journal Systems (OJS) [21]. Наборы метаданных теперь формируются автоматически с помощью программных инструментов, разработанных в редакции (<http://ojs.kpfu.ru/index.php/elbib>). Архив статей, опубликованных, начиная с 2015 года, был выбран для подготовки к индексации в DBLP. Необходимыми метаданными являются: идентификатор публикации, фамилии и имена авторов, название работы, год издания, том, номер, начальная и конечная страницы статьи в номере журнала и URL полного текста статьи.

Нормализация в формат DBLP происходит в три этапа: экстракция требуемых метаданных, дополнение метаданных и их нормализация в формат DBLP.

С помощью программы, разработанной на языке C#, и средств расширения System:XML выполняется последовательная обработка файлов коллекции и, как результат, формируется набор метаданных каждого документа. На следующем этапе метаданные пополняются информацией о статье и её авторах на английском языке. Эта информация импортируется с англоязычной версии сайта журнала средствами расширения HTMLAgilityPack. Поскольку англоязычная информация об авторах неполная – указаны только фамилии и инициалы – производится транслитерация имен с русскоязычной страницы. Результатом этой работы стало включение журнала “Russian Digital Libraries Journal” и статей, опубликованных в нем в 2015–2018 годах, в базу DBLP (<https://dblp.uni-trier.de/db/journals/rdlj/>).

4. Цифровое хранилище на основе DSPACE

Одной из важных задач при работе с цифровыми математическими библиотеками является автоматизированная интеграция репозитория математических документов в другие информационные системы. Данный процесс основан на модели агрегирования и распространений метаданных. Эта модель OAI Protocol for Metadata Harvesting (далее OAI-PMH) [22] поддерживается большинством систем, предназначенных для хранения информационных ресурсов. Данную систему поддерживают, например, такие цифровые библиотеки, как EuDML, NumDam, DBLP.

Инициатива «Открытые архивы» (OAI) разрабатывает и продвигает стандарты интероперабельности с целью эффективного распространения электронных ресурсов, а также повышения доступности обмена научной информацией. Протокол OAI-PMH требует обязательного включения в описание ресурса набора метаданных Дублинского ядра (Dublin Core, <http://dublincore.org/>). Применение протокола OAI-PMH требует обмена информацией в рамках жестко установленных схем данных. В конкретных информационных системах такие схемы, как правило, не реализованы, поэтому необходимы динамическая конвертация метаданных или автоматическая подготовка метаданных в формате, пригодном для OAI-PMH.

Кроме того, для организации работы с ОАИ-РМН необходимо использовать систему поддержки цифрового хранилища. Обзор различных цифровых хранилищ приведен в [23]. Наиболее популярными из них являются DSpace, Eprints, Fedora и Greenstone. Мы используем систему Dspace – приложение с открытым исходным кодом (лицензия BSD), которое является кроссплатформенным и основано на Java. Для хранения метаданных применены СУБД Oracle или PostgreSQL. Для базовой организации данных зафиксирована модель данных, основанная на схеме Dublin Core. Также имеется возможность загружать свои форматы метаданных и конвертеры, что делает данную систему наиболее привлекательной для использования. Таким образом, реализована возможность автоматической конвертации различных форматов данных в Dublin Core, что позволяет осуществлять харвестинг метаданных. Однако для загрузки пользовательских форматов необходимо разрабатывать специализированные системы конвертации метаданных.

Заключение

С целью интеграции электронных математических коллекций Казанского университета в международное научное пространство разработаны алгоритмы формирования метаданных этих коллекций и документов, входящих в них, в соответствии с форматами цифровых математических библиотек и наукометрических баз данных. Представлены методы нормализации метаданных электронных математических коллекций в соответствии с XML-схемами NISO JATS и DBLP.

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 научного проекта № 18-47-160012. Настоящая статья содержит также результаты проекта «Разработка технологий управления математическими знаниями на основе цифровой математической библиотеки Lobachevskii-DML», выполняемого в рамках реализации Программы Центра компетенций Национальной технологической инициативы «Центр хранения и анализа больших данных», поддерживаемого Министерством науки и высшего образования Российской Федерации по Договору МГУ им. М.В. Ломоносова с Фондом поддержки проектов Национальной технологической инициативы от 11.12.2018 № 13/1251/2018.

Литература

1. Developing a 21st Century Global Library for Mathematics Research. Washington: The National Academies Press. — 2014. — 131 p.
2. Ion P.D.F., Watt S.M. The Global Digital Mathematics Library and the International Mathematical Knowledge Trust // ICM 2017: Intelligent Computer

- Mathematics, 2017. Lecture Notes in Artificial Intelligence. Springer. — 2017. — Vol. 10383. — P. 56–69. — https://doi.org/10.1007/978-3-319-62075-6_5.
3. Bouche T. Reviving the free public scientific library in the digital age? the EuDML project // In: Kaiser K., Krantz S.G., Wegner B. (Eds.) Topics and Issues in Electronic Publishing JMM/AMS Special Session. FIZ Karlsruhe, 2013. — P. 57–80. — URL: <https://www.emis.de/proceedings/TIEP2013/05bouche.pdf>.
 4. Bouche T., Rákosník J. Report on the EuDML External Cooperation Model // In: Kaiser K., Krantz S.G., Wegner B. (Eds.) Topics and Issues in Electronic Publishing, JMM, Special Session, San Diego, 2013. — P. 99–108. — URL: https://www.emis.de/proceedings/TIEP2013/07bouche_rakosnik.pdf.
 5. Chebukov D.E., Izaak A.D., Misyurina O.G., Pupyrev Yu.A., 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. — Vol. 7961. — P. 344–348. — https://doi.org/10.1007/978-3-642-39320-4_26.
 6. Elizarov A.M., Lipachev E.K., Zuev D.S. Digital Mathematical Libraries: Overview of Implementations and Content Management Services // CEUR Workshop Proceedings. — 2017. — Vol. 2022. — P. 317–325.
 7. Bartošek M., Rákosník J. DML-CZ: The Experience of a Medium-Sized Digital Mathematics Library // Notices of the AMS. — 2013. — Vol. 60. — No 8. — P. 1028–1033. — <http://dx.doi.org/10.1090/noti1031>.
 8. Elizarov A.M., Kirillovich A.V., Lipachev E.K., Nevzorova O.A. Mathematical Knowledge Management: Ontological Models and Digital Technology // CEUR Workshop Proceedings. — 2016. — Vol. 1752. — P. 44–50.
 9. Elizarov A.M., Lipachev E.K. Lobachevskii DML: Towards a Semantic Digital Mathematical Library of Kazan University // CEUR Workshop Proceedings. — 2017. — Vol. 2022. — P. 326–333.
 10. 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. Москва: ИПМ им. М. В. Келдыша, 2017. — С. 130–136. — <https://doi.org/10.20948/abrau-2017-73>.
 11. Bouche T., Labbe O. The New Numdam Platform // CICM 2017: Intelligent Computer Mathematics, 2017. — P. 70–82. — https://doi.org/10.1007/978-3-319-62075-6_6.
 12. Elizarov A.M., Lipachev E.K., Khaidarov Sh.M. Automated Processing Service System of Large Collections of Scientific Documents // CEUR Workshop Proceedings. — 2016. — Vol. 1752. — P. 58–64.
 13. Elizarov A.M., Khaydarov Sh.M., Lipachev E.K. Scientific Documents Ontologies for Semantic Representation of Digital Libraries // Proc. of the 2nd Russia and Pacific Conf. on Computer Technology and Applications. — 2017. — P. 1–5. — <https://doi.org/10.1109/RPC.2017.8168064>.
 14. Elizarov A.M., Kirillovich A.V., Lipachev E.K., Nevzorova O.A. Digital Ecosystem OntoMath: Mathematical Knowledge Analytics and Management //

- Communications in Computer and Information Science, Springer, 2017. — Vol. 70. — P. 33–46. — https://doi.org/10.1007/978-3-319-57135-5_3.
15. “ANSI/NISO Z39.96-2019, JATS: Journal Article Tag Suite”. National Information Standards Organization. 8 February 2019. — 652 p. — URL: https://groups.niso.org/apps/group_public/download.php/21030/ANSI-NISO-Z39.96-2019.pdf.
 16. Елизаров А.М., Зайцева Н.В., Зуев Д.С., Липачёв Е.К., Хайдаров Ш.М. Сервисы формирования метаданных цифровых документов в форматах международных наукометрических баз данных // Научный сервис в сети Интернет: труды XX Всероссийской научной конференции (17–22 сентября 2018 г., г. Новороссийск). — М.: ИПМ им. М.В. Келдыша, 2018. — С. 175–185. — URL: <http://keldysh.ru/abrau/2018/theses/53.pdf>. — <https://doi.org/10.20948/abrau-2018-53>.
 17. Jost M., Bouche T., Goutorbe C., Jorda J.P. D3.2: The EuDML metadata schema. — URL: <http://www.mathdoc.fr/publis/d3.2-v1.6.pdf>.
 18. EuDML metadata schema specification (v2.0–final). — <https://initiative.eudml.org/eudml-metadata-schema-specification-v20-final>.
 19. Герасимов А.Н., Елизаров А.М., Липачев Е.К. Формирование метаданных для международных баз цитирования в системе управления электронными научными журналами // Электронные библиотеки. — 2015.— Т. 18. — № 1-2. — С. 6-31.
 20. Герасимов А.Н., Елизаров А.М., Липачев Е.К., Хайдаров Ш.М. Методы автоматизированного извлечения метаданных научных публикаций для библиографических и реферативных баз цитирования // Сб. науч. статей XIX Объединенной конф. "Интернет и современное общество" IMS-2016. — 2016. — С. 41-48. — URL: <https://openbooks.itmo.ru/ru/file/4086/4086.pdf>.
 21. Ахметов Д.Ю., Елизаров А.М., Липачёв Е.К. Сервис-ориентированная информационная система научного журнала «Электронные библиотеки» // Электронные библиотеки. — 2016. — Т. 19. — № 1. — С. 2–39.
 22. Open Archives Initiative Protocol for Metadata Harvesting. — URL: <http://www.openarchives.org/OAI/openarchivesprotocol.html>.
 23. Федотов А.М., Байдавлетов А.Т., Жижимов О.Л., Самбетбаева М.А., Федотова О.А. Цифровой репозиторий в научно-образовательной информационной системе // Вестн. Новосиб. гос. ун-та. Серия: Информационные технологии. — 2015. — Т. 13, вып. 3. — С. 68–86.

Reference

1. Developing a 21st Century Global Library for Mathematics Research. Washington: The National Academies Press. — 2014. — 131 p.
2. Ion P.D.F., Watt S.M. The Global Digital Mathematics Library and the International Mathematical Knowledge Trust // ICM 2017: Intelligent Computer Math-

- ematics, 2017. Lecture Notes in Artificial Intelligence. Springer. — 2017. — Vol. 10383. — P. 56–69. — https://doi.org/10.1007/978-3-319-62075-6_5.
3. Bouche T. Reviving the free public scientific library in the digital age? the EuDML project // In: Kaiser K., Krantz S.G., Wegner B. (Eds.) Topics and Issues in Electronic Publishing JMM/AMS Special Session. FIZ Karlsruhe, 2013. — P. 57–80. — URL: <https://www.emis.de/proceedings/TIEP2013/05bouche.pdf>.
 4. Bouche T., Rákosník J. Report on the EuDML External Cooperation Model // In: Kaiser K., Krantz S.G., Wegner B. (Eds.) Topics and Issues in Electronic Publishing, JMM, Special Session, San Diego, 2013. — P. 99–108. — URL: https://www.emis.de/proceedings/TIEP2013/07bouche_rakosnik.pdf.
 5. 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. — Vol. 7961. — P. 344–348. — https://doi.org/10.1007/978-3-642-39320-4_26.
 6. Elizarov A.M., Lipachev E.K., Zuev D.S. Digital Mathematical Libraries: Overview of Implementations and Content Management Services // CEUR Workshop Proceedings. — 2017. — Vol. 2022. — P. 317–325.
 7. Bartošek M., Rákosník J. DML-CZ: The Experience of a Medium-Sized Digital Mathematics Library // Notices of the AMS. — 2013. — Vol. 60. — No 8. — P. 1028–1033. — <http://dx.doi.org/10.1090/noti1031>.
 8. Elizarov A.M., Kirillovich A.V., Lipachev E.K., Nevzorova O.A. Mathematical Knowledge Management: Ontological Models and Digital Technology // CEUR Workshop Proceedings. — 2016. — Vol. 1752. — P. 44–50.
 9. Elizarov A.M., Lipachev E.K. Lobachevskii DML: Towards a Semantic Digital Mathematical Library of Kazan University // CEUR Workshop Proceedings. — 2017. — Vol. 2022. — P. 326–333.
 10. Elizarov A.M., Lipachev E.K. Semanticheskie metody i instrumenty jelektronnoj matematicheskoy biblioteki Lobachevskii-DML // Nauchnyj servis v seti Internet: trudy XIX Vserossijskoj nauchnoj konferencii. Moskva: IPM im. M. V. Keldysha, 2017. — S. 130–136. — <https://doi.org/10.20948/abrau-2017-73>.
 11. Bouche T., Labbe O. The New Numdam Platform // CICM 2017: Intelligent Computer Mathematics, 2017. — P. 70–82. — https://doi.org/10.1007/978-3-319-62075-6_6.
 12. Elizarov A.M., Lipachev E.K., and Khaidarov Sh.M. Automated Processing Service System of Large Collections of Scientific Documents // CEUR Workshop Proceedings. — 2016. — Vol. 1752. — P. 58–64.
 13. Elizarov A.M., Khaydarov Sh.M., Lipachev E.K. Scientific Documents Ontologies for Semantic Representation of Digital Libraries // Proc. of the 2nd Russia

- and Pacific Conf. on Computer Technology and Applications. — 2017. — P. 1–5. — <https://doi.org/10.1109/RPC.2017.8168064>.
14. Elizarov A.M., Kirillovich A.V., Lipachev E.K., Nevzorova O.A. Digital Ecosystem OntoMath: Mathematical Knowledge Analytics and Management // Communications in Computer and Information Science, Springer, 2017. — Vol. 70. — P. 33-46. — https://doi.org/10.1007/978-3-319-57135-5_3.
 15. “ANSI/NISO Z39.96-2019, JATS: Journal Article Tag Suite”. National Information Standards Organization. 8 February 2019. — 652 p. — URL: https://groups.niso.org/apps/group_public/download.php/21030/ANSI-NISO-Z39.96-2019.pdf.
 16. Elizarov A.M., Zaitseva N.V., Zuev D.S., Lipachev E.K., Khaidarov S.M. Services for Formation of Digital Documents Metadata in the Formats of International Science-based Databases // Nauchnyj servis v seti Internet: trudy XX Vserossijskoj nauchnoj konferencii. Moskva: IPM im. M. V. Keldysha, 2018. — S. 175–185. — <https://doi.org/10.20948/abrau-2018-53>.
 17. Jost M., Bouche T., Goutorbe C., Jorda J.P. D3.2: The EuDML metadata schema. — URL: <http://www.mathdoc.fr/publis/d3.2-v1.6.pdf>.
 18. EuDML metadata schema specification (v2.0–final). — <https://initiative.eudml.org/eudml-metadata-schema-specification-v20-final>.
 19. Gerasimov A.N., Elizarov A.M., Lipachev E.K. Subsystem of formation metadata for science index databases on management platform electronic scientific journals // Russian Digital Libraries Journal. — 2015. — V. 18. — No. 1-2. — P. 6-31.
 20. Gerasimov A.N., Elizarov A.M., Lipachev E.K., Khaydarov S.M. Automated methods of metadata extraction from scientific publications for bibliographic databases // Internet and Modern Society (IMS-2016) — 2016. — P. 41-48. — URL: <https://openbooks.itmo.ru/ru/file/4086/4086.pdf>.
 21. Akhmetov D., Elizarov A., Lipachev E. Service-oriented information system of "Russian Digital Libraries Journal" // Russian Digital Libraries Journal. — 2016. — V. 19, No 1. — S. 2–39.
 22. Open Archives Initiative Protocol for Metadata Harvesting. — URL: <http://www.openarchives.org/OAI/openarchivesprotocol.html>.
 23. Fedotov A.M., Baidavletov A.T., Zhizhimov O.L., Sambetbayeva M.A., Fedotova O. A. Digital Repository of Scientific and Educational Information System // Vestn. Novosib. gos. un-ta. Serija: Informacionnye tehnologii. — 2015. — T. 13, vyp. 3. — S. 68–86.