



А.М. Елизаров, Н.В. Зайцева, Д.С. Зуев,
Е.К. Липачёв, Ш.М. Хайдаров

**Сервисы формирования метаданных
цифровых документов в форматах
международных наукометрических
баз данных**

Рекомендуемая форма библиографической ссылки

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> doi:[10.20948/abrau-2018-53](https://doi.org/10.20948/abrau-2018-53)

Размещена также [презентация к докладу](#)

Сервисы формирования метаданных цифровых документов в форматах международных наукометрических баз данных

А.М. Елизаров^{1,2}, Н.В. Зайцева², Д.С. Зуев¹, Е.К. Липачёв², Ш.М. Хайдаров²

¹Высшая школа информационных технологий и информационных систем,

²Институт математики и механики им. Н.И. Лобачевского
Казанского (Приволжского) федерального университета

Аннотация. Приведен обзор современных наукометрических баз данных и освещена специфика представления в них научных материалов. Представлены методы интеграции цифровых научных коллекций, основанные на автоматизации процесса создания метаданных для документов, входящих в эти коллекции. Отмечены особенности формирования метаданных международных наукометрических баз данных по математическим и компьютерным наукам. Приведен алгоритм автоматизированного формирования метаданных в формате Российского индекса научного цитирования (РИНЦ). Для автоматического разбора текста статей создано несколько шаблонов регулярных выражений, с помощью которых выделялись основные группы метаданных (список авторов, номер тома, страницы и т. д.). Алгоритм реализован в виде сервиса, состоящего из модулей анализа структуры документов, автоматического выбора документов согласно установленному порядку (например, лексикографическому), извлечения блока аннотации, модулей генерации алфавитного указателя, формирования библиографического описания статьи для записи в блок колонтитулов этой статьи, конвертации документов в pdf-формат в соответствии с установленными параметрами. Завершающим является модуль формирования метаданных для экспорта в РИНЦ. Алгоритм апробирован на коллекции статей журнала «Электронные библиотеки». Представлен сервис формирования метаданных документов цифровой коллекции Lobachevskii DML в соответствии со схемами фундаментальных метаданных The European Digital Mathematical Library (EuDML) и библиографической базы данных DBLP. Подготовлены шаблоны представления метаданных статей цифровых коллекций Lobachevskii DML в соответствии со схемой NISO JATS V1.0. Разработаны плагины для системы Open Journal System, позволяющие для загружаемых статей генерировать метаданные для наукометрических баз данных.

Ключевые слова: метаданные, цифровые коллекции, семантические методы, наукометрические базы данных, математические и компьютерные науки, структурный и стилиевой анализ цифровых документов, Европейская цифровая математическая библиотека EuDML, библиографическая база данных DBLP.

Services for Formation of Digital Documents Metadata in the Formats of International Science-based Databases

A.M. Elizarov^{1,2}, N.V. Zaitseva², D.S. Zuev¹, E.K. Lipachev², S.M. Khaidarov²

¹ *Higher School of Information Technologies and Intelligent Systems,*

² *N.I. Lobachevskii Institute of Mathematics and Mechanics*

Kazan (Volga Region) Federal University

Abstract. This paper contains the review of modern scientometric databases. The specificity of the representation of scientific materials in them is highlighted. The integration methods based on automation of the process of creating metadata for documents, included in digital scientific collections, are presented. Features of the formation of metadata for international scientometric databases on mathematical and computer sciences are noted. An algorithm for the automated formation of metadata in the format of the Russian scientific citation index (RSCI) is given. To automatically parse the text of articles, several regular expression patterns have been created, with the help of which the main metadata groups were selected. The algorithm is implemented as a service, consisting of modules for analyzing the structure of documents, automatically selecting documents according to the established order (for example, lexicographic), extracting the annotation block, the alphabetical index generating module, creating a bibliographic description of the article for writing headers of this article, converting documents to the portable document format (pdf), according to the determined parameters. The final module is the formation of metadata for exports to the RSCI. Approbation of the algorithm for the collection of articles of the journal "Russian Digital Libraries" was noted. The service for the formation of metadata for the documents of the digital collection Lobachevskii DML, made in accordance with the diagrams of the fundamental metadata of the European Digital Mathematical Library (EuDML) and the bibliographic database DBLP, is presented. Templates for showing metadata of articles of digital collections Lobachevskii DML in accordance with the scheme NISO JATS V1.0 are prepared. Plugins for the Open Journal System, allowing generation of metadata for science-based databases for downloadable articles are developed.

Keywords: metadata, digital collections, semantic methods, scientometric databases, mathematical and computer sciences, structural and stylistic analysis of digital documents, European Digital Mathematical Library, EuDML, DBLP.

1. Современные наукометрические базы данных

В настоящее время определяющую роль в процессах анализа научной информации и дальнейшего ее распространения в Сети играют международные и региональные наукометрические базы данных, общее количество которых весьма велико. Например, на сайте <http://dnrpress.ru/международные-наукометрические-базы/> представлен список из более чем 220 баз данных международного и регионального, реферативного и полнотекстового, мультимедийного

циплинарного и специализированного характера, а также перечень цифровых библиотек наиболее известных научно-образовательных учреждений.

Самой авторитетной в мире и исторически первой аналитической и цитатной базой данных журнальных статей (создана в 1965 году) является информационная система Web of Science (WoS, <http://apps.webofknowledge.com/>), которая размещена на поисковой платформе Web of Knowledge (<http://apps.webofknowledge.com/>) и охватывает материалы по естественным, техническим, биологическим, общественным, гуманитарным наукам и искусству. Эта платформа объединяет реферативные базы данных публикаций в научных журналах и патентов, в том числе базы, учитывающие взаимное цитирование публикаций, обладает встроенными возможностями поиска, анализа и управления библиографической информацией и развивается сегодня компанией Clarivate Analytics (<https://clarivate.com/products/web-of-science/>).

Крупнейшей в мире мультидисциплинарной библиографической, реферативной и цитатной базой данных является информационная система Scopus (SciVerse Scopus) (<http://www.scopus.com>), созданная в 2005 году издательской корпорацией Elsevier. Scopus охватывает свыше 18 тыс. научных журналов мира (включая около 200 российских), около 70 млн. публикаций (с ежегодным приростом около 2,8 млн.), более 13 млн. патентов и материалы научных конференций. Scopus в отличие от WoS гораздо шире отражает естественные науки и технику и не включает издания по гуманитарным дисциплинам и искусству.

Российский индекс научного цитирования (РИНЦ, https://elibrary.ru/project_risc.asp), созданный в 2006 году, – это российская национальная информационно-аналитическая система, аккумулирующая сегодня около 12 млн. публикаций российских авторов, а также информацию о цитировании этих публикаций из более чем 6 тыс. российских журналов. Она содержит эффективные механизмы оценки качества научных журналов и является мощным аналитическим инструментом, позволяющим осуществлять оценку результативности и эффективности деятельности научно-исследовательских организаций и непосредственно ученых, а также уровень научных журналов.

Отметим также, что существуют международные наукометрические базы данных для конкретных предметных областей. Примерами служат Международная информационная система по сельскохозяйственным наукам и технологиям AGRIS (International Information System for the Agricultural Sciences and Technology, <http://agris.fao.org/>) и библиографическая база данных GeoRef (<http://www.agiweb.org/>) по геологии и геонаукам.

Так как результаты наших исследований, описанные ниже, связаны с публикациями по математическим и компьютерным наукам, назовем некоторые известные наукометрические базы данных в этой предметной области.

Современной информационной системой, предоставляющей российским и зарубежным математикам различные возможности в поиске и обработке математической информации, является Общероссийский математический портал Math-Net.Ru (<http://www.mathnet.ru/>).

Европейский проект (The European Digital Mathematical Library, EuDML, <https://initiative.eudml.org/>) нацелен на построение цифровой библиотеки, интегрирующей математический контент, и организацию доступа к нему с единой платформы.

DBLP (<https://dblp.uni-trier.de>) – это онлайн библиографическая база данных публикаций в области компьютерных наук, создаваемая в Германии совместно University of Trier и Schloss Dagstuhl. По мнению ее создателей, миссия DBLP заключается в том, чтобы поддерживать ученых в области компьютерных наук, предоставляя им бесплатный доступ к высококачественным библиографическим метаданным и ссылкам на электронные издания и цифровые публикации. На текущий момент времени DBLP индексирует более 4,1 млн. публикаций более чем 1,7 млн. авторов.

Итак, в настоящее время имеется большое количество наукометрических баз данных, в которых могут быть размещены и проиндексированы (причем, одновременно в нескольких) цифровые документы различных научных коллекций (журналов, сборников материалов научных конференций, препринтов и др.). Основным элементом описания этих документов при загрузке в названные базы данных являются метаданные, сформированные в соответствии с форматами, принятыми в соответствующих информационных системах. Как правило, эти форматы отличаются. Например, собственную схему метаданных имеет DBLP, она описана на официальном сайте проекта, где также приведены образцы оформления метаданных статей и сборников трудов конференций. Далее, формирование метаданных и загрузка соответствующих документов в наукометрические базы данных, естественно, должны быть автоматизированы. Следовательно, актуальна задача автоматизации формирования метаданных цифровых документов в форматах соответствующих наукометрических баз данных. Результаты, полученные нами в этом направлении, описаны ниже.

2. Автоматизированное формирование метаданных в формате РИНЦ

Для загрузки метаданных отдельных статей, выпусков журналов и книг в базу данных РИНЦ, как правило, используется онлайн-программа разметки Artculus (<http://elibrary.ru/projects/contracts/publisher/messages/messages.asp>). Подготовка метаданных в этой программе требует большого объема ручной работы. Нами предложен алгоритм автоматизации формирования блока метаданных, который представлен в [1, 2]. Апробация алгоритма выполнялась в несколько этапов на массивах естественнонаучных журналов, в частности, на архиве статей журнала «Электронные библиотеки» (Russian Digital Libraries Journal) за 1998–2014 годы, который в результате полностью загружен в РИНЦ. Опишем примененные методы.

Для извлечения метаданных автоматически производился разбор файла статьи. Статьи в формате .doc были предварительно преобразованы в формат .docx. С использованием структуры .docx-файла извлекался основной текст документа, сведения об авторах и их аффилиации, название работы, аннотация и

библиографические данные. В качестве отличительных признаков блока библиографии алгоритм использует наличие библиографических записей на русском и английском языках, заголовка раздела и следование после фамилии инициалов авторов. Для автоматического разбора библиографии создано несколько шаблонов регулярных выражений, с помощью которых выделялись основные группы метаданных (список авторов, номер тома, страницы и т. д.). Далее проводился разбор каждой группы, в частности, осуществлялось выделение каждого автора из списка авторов статьи. Основными модулями системы являются:

- модуль извлечения метаданных из документов коллекции на основе анализа их структуры и форматов представления информации;
- модуль автоматического выбора документов согласно установленному порядку (например, лексикографическому или по спискам авторов);
- модуль извлечения блока аннотации из документов коллекции;
- модуль формирования алфавитного указателя;
- модуль автоматического формирования библиографического описания каждой статьи коллекции с последующей записью этой информации в блок колонтитулов обрабатываемого документа;
- модуль конвертации документов в pdf-формат в соответствии с установленными параметрами;
- модуль подготовки метаданных для экспорта в РИНЦ.

Разработанный алгоритм реализован в виде сервиса [3] и в настоящее время используется в журнале «Электронные библиотеки» (<http://ojs.kpfu.ru/index.php/elbib>). Также этот сервис был успешно использован при обработке больших массивов научных документов [4]. Мы планируем улучшить алгоритм, добавив в него возможность программного построения регулярных выражений на основе шаблона представленных материалов.

3. Сервис преобразования метаданных в формат EuDML

Проект EuDML [5–7] ставит основными целями разработку и построение совместной цифровой библиотеки, интегрирующей математический контент, представленный его партнерами. Этот проект можно рассматривать как один из этапов построения Всемирной цифровой математической библиотеки (см., например, [8–10]).

Проект EuDML – это агрегатор математических ресурсов и их метаданных, который поддерживается Европейским математическим обществом (European Mathematical Society). EuDML нацелен на сбор и распространение максимального объема математических ресурсов для повышения их доступности и предоставления к ним открытого доступа, организации этих ресурсов в форме цифровых математических коллекций, обладающих различными поисковыми и иными сервисами, и является первой попыткой крупномасштабной реализации цифровой математической библиотеки. Фактически этот проект разрабатывается и поддерживается сетью научных организаций Евросоюза [9]. Информаци-

онная система, реализуемая в рамках проекта, состоит из репозитория метаданных, поискового механизма, модулей работы с метаданными, «анализатора ассоциаций» и функций аннотирования и предоставления доступа.

В [6] описана схема метаданных, принятая в EuDML. Они разделены на основные, фундаментальные и дополнительные (см. <https://initiative.eudml.org/eudml-metadata-schema-specification-v20-final>). Репозиторий метаданных является ядром всей системы. Он работает со сборщиком метаданных по протоколу OAI-PMH и конвертирует метаданные во внутреннюю схему EuDML.

Работа с метаданными обеспечена набором программных инструментов, каждый из которых позволяет расширять или дополнять метаданные существующих информационных ресурсов. Например, это могут быть добавления ключевых слов или многоязычных метаданных путем слияния информации из разных баз данных или генерация MathML-описаний для математических выражений.

Модуль «Анализатор ассоциаций» находит и фиксирует отношения между отдельными сущностями или ресурсами. Функция аннотирования предоставляет механизмы для добавления и редактирования новой информации к отдельным элементам в репозиториях. Компонент управления доступом обеспечивает поддержку и управление политиками доступа к элементам репозитория.

Наконец, пользовательские и программные интерфейсы системы обеспечивают доступ к собранным ресурсам на разных уровнях. EuDML поддерживает ряд протоколов для межсистемной интеграции: OAI-PMH, REST API, OpenSearch и JSON.

EuDML стремится быть открытым источником надежных математических знаний, для этих целей выработаны определенные политики и правила размещения и доступности индексируемых ресурсов: все тексты статей должны быть научно обоснованы и официально опубликованы; все ресурсы должны быть в открытом доступе; полный текст информационного ресурса должен физически храниться в репозитории одного из учреждений-партнеров EuDML.

Для журналов «Lobachevskii Journal of Mathematics», «Известия вузов. Математика» и «Электронные библиотеки» нами подготовлены шаблоны представления метаданных в соответствии со схемой Journal Archiving and Interchange Tag Suite (NISO JATS V1.0), принятой EuDML для описания журнальных статей. При проведении метаописания коллекции статей журнала «Электронные библиотеки» до 2015 года пришлось решить задачу включения транслитерированных строк с названиями статей, а также фамилиями авторов. При этом учтено, что имеется несколько систем транслитерации. Для статей, поступающих в журналы через систему Open Journal System, указанные метаданные формируются автоматически с помощью разработанного нами плагина. Также при метаописании статей журнала «Известия вузов. Математика» учтено, что статьи с 1974 года переводятся на английский язык и выходят с другим библиографическим описанием в журнале «Russian Mathematics» (до 1991 года – Soviet Mathematics).

4. Сервис автоматизированного формирования метаданных для DBLP

Многие российские ученые, областью научных интересов которых являются информатика, информационные и компьютерные технологии, сталкиваются с тем, что их работы лишь в малой доле индексируются в Scopus и в гораздо меньшей степени представлены в WoS. Такая ситуация обусловлена тем, что многие страны с развитой экономикой, к числу которых относится и Россия, стремятся к накоплению и обмену научными знаниями на родном языке, тогда как базы данных WoS и Scopus существенно большее внимание уделяют англоязычным работам. Кроме того, журналы по информатике, проиндексированные в WoS и Scopus, составляют лишь незначительную часть всех журналов, проиндексированных в этих базах данных. Сложившаяся ситуация способствовала созданию специализированных информационных систем, ориентированных на индексацию и учет цитирования публикаций по компьютерным наукам. Одной из них стала библиографическая база данных DBLP, созданная в 1990-х годах учеными университетов Германии для отслеживания публикаций по системам баз данных Data Base и логическому программированию Logic Programming, но утратившая в дальнейшем свое первоначальное предназначение. Сегодня аббревиатура DBLP понимается как «Digital Bibliography & Library Project» и олицетворяет веб-сервис, оперирующий материалами ведущих журналов и крупных конференций по всем отраслям компьютерных наук, которые размещены на трех сайтах (<http://www.informatik.uni-trier.de/~ley/db>, <http://dblp.uni-trier.de/db/>, <http://dblp.dagstuhl.de/db/>), синхронизированных между собой. База данных открытая, любой ресурс по тематике компьютерных наук может быть проиндексирован ею. Вместе с тем, для этого необходимо выполнение ряда условий, таких, как наличие полных метаданных и рефератов работ в открытом доступе, doi или возможность доступа к индексируемым ресурсам поисковыми роботам.

Для каждой публикации из любой коллекции DBLP метаданные хранятся в отдельном файле, созданном с помощью языка-разметки html, и образуют следующий перечень, являющийся, как нам представляется, далеко не полным: имена авторов, название публикации, название журнала, том, номер, год издания, количество страниц публикации, url-адрес.

Сервис DBLP имеет, по нашему мнению, тот недостаток, что применяемые принципы отбора и обработки библиографической информации не совсем ясны. Далее, пока не существует общедоступной html-формы для ввода библиографических данных, поэтому отдельные авторы не могут самостоятельно разместить свои публикации в этой базе данных. Кроме того, DBLP размещает метаданные документов отдельных выпусков журналов, полученных от издателя или организаторов конференции, только по предварительному договору о сотрудничестве. При этом индексирование публикаций занимает достаточно большой промежуток времени – от нескольких месяцев до года. Появление в DBLP новых журналов по соответствующей тематике или материалов серии

конференций возможно только по решению координаторов веб-сайта после соответствующего обращения к ним.

Заключение

Для автоматизированного формирования метаданных для библиографических баз данных, описанных выше, нами адаптированы алгоритмы выделения метаданных из корпуса физико-математических текстов, предложенные в работах [1–3, 11]. Так, с использованием сервиса [11] формируются обязательные метаданные, в соответствии со схемой EuDML [12], а также в формате DBLP [13] для последующей загрузки. Этот сервис позволяет автоматически проводить обработку больших коллекций физико-математических документов, включая валидацию документов и их преобразование в соответствии с правилами формирования коллекций, семантический анализ документов, извлечение метаданных и др. (см. [14, 15]).

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.

Литература

1. Герасимов А.Н., Елизаров А.М., Липачёв Е.К. Формирование метаданных для международных баз цитирования в системе управления электронными научными журналами // Электронные библиотеки. – 2015. – №1-2 (18). – С. 6–31.
2. 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 конференции «Интернет и современное общество» IMS-2016. – СПб, 2016. – С. 41–48.
3. Свидетельство о государственной регистрации программы для ЭВМ. Программа автоматизированного формирования метаданных в формате Российского индекса научного цитирования для статей журнала «Электронные библиотеки» / А.М. Елизаров, Е.К. Липачёв, Ш.М. Хайдаров; заявитель и правообладатель ФГАОУ ВО КФУ (RU). – №2017663206; заявл. 19.12.2017; опубл. 16.02.2018, Реестр программ для ЭВМ.
4. Elizarov A.M., Lipachev E.K., and Khaydarov S.M. Automated System of Services for Processing of Large Collections of Scientific Documents // CEUR Workshop Proceedings. – 2016. – Vol. 1752. – P. 58–64.
5. Sylwestrzak W., Borbinha J., Bouche T., Nowiński A., and Sojka P. EuDML – Towards the European Digital Mathematics Library // In: Sojka P. (ed.): Towards a Digital Mathematics Library. Paris, France, July 7–8th, 2010. Masaryk

- University Press, Brno, Czech Republic, 2010. – P. 11–26. URL: <https://dml.cz/handle/10338.dmlcz/702569>.
6. Bouche T., Goutorbe C., Jorda J.-P., and Jost M. The EuDML Metadata Schema: Version 1.0 // In: Sojka P., Bouche T. (eds.): Towards a Digital Mathematics Library. Bertinoro, Italy, July 20–21st, 2011. Masaryk University Press, Brno, Czech Republic, 2011. P. 45–61. URL: <https://dml.cz/handle/10338.dmlcz/702602>.
 7. Bouche T., and Rákosník J. Report on the EuDML External Cooperation Model // In: Kaiser K., Krantz S., Wegner B. (Eds.): Topics and Issues in Electronic Publishing, JMM, Special Session, San Diego, January 2013. – P. 99–108. URL: https://www.emis.de/proceedings/TIEP2013/07bouche_rakosnik.pdf.
 8. Developing a 21st Century Global Library for Mathematics Research. Washington, D.C.: The National Academies Press, Washington, D.C (2014). arxiv.org/pdf/1404.1905, <https://www.nap.edu/catalog/18619/developing-a-21st-century-global-library-for-mathematics-research>
 9. Elizarov A.M., Lipachev E.K., and Zuev D.S. Digital Mathematical Libraries: Overview of Implementations and Content Management Services // CEUR Workshop Proceedings. – 2017. – Vol. 2022. – P. 317–325.
 10. Ion P.D.F., and Watt S.M. The Global Digital Mathematics Library and the International Mathematical Knowledge Trust // Geuvers H. et al. (Eds.): CICM 2017, LNAI 10383. 2017. – P. 56–69. https://doi.org/10.1007/978-3-319-62075-6_5.
 11. Свидетельство о государственной регистрации программы для ЭВМ. Программный комплекс выделения метаданных из коллекций физико-математических документов, представленных в формате OpenXML / А.М. Елизаров, Е.К. Липачёв, Ш.М. Хайдаров; заявитель и правообладатель ФГАОУ ВО КФУ (RU). – №2016616511; заявл. 21.06.16; опубли. 08.08.16, Реестр программ для ЭВМ.
 12. EuDML metadata schema specification (v2.0 – final). URL: <https://initiative.eudml.org/eudml-metadata-schema-specification-v20-final>.
 13. DBLP XML submission format guidelines. URL: <http://dblp.uni-trier.de/faq/dblpsubmission.xsd>.
 14. 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>. <https://doi.org/10.20948/abrau-2017-73>.
 15. Elizarov A.M., and Lipachev E.K. Lobachevskii DML: Towards a Semantic Digital Mathematical Library of Kazan University // CEUR Workshop Proceedings. – 2017. – Vol. 2022. – P. 326–333.

References

1. Gerasimov A.N., Elizarov A.M., and Lipachev E.K. Subsystem of Formation Metadata for Science Index Databases on Management Platform Electronic Scientific Journals // Russian Digital Libraries Journal. – 2015. – No 1-2 (18). – P. 6–31.
2. Gerasimov A.N., Elizarov A.M., Lipachev E.K., and Khaydarov S.M. Metody avtomati-zirovannogo izvlecheniya metadannykh nauchnykh publikatsiy dlya bibliografiche-skikh i referativnykh baz tsitirovaniya // Sb. nauch. statey XIX Ob"yedinennoy konferentsii «Internet i sovremennoye obshchestvo» IMS-2016. – SPb, 2016. – P. 41–48.
3. Svidetel'stvo o gosudarstvennoy registratsii programmy dlya EVM. Programma avtomatizirovannogo formirovaniya metadannykh v formate Rossiyskogo indeksa nauchnogo tsitirovaniya dlya statey zhurnala «Elektronnyye biblioteki» / A.M. Elizarov, E.K. Lipachev, S.M. Khaydarov; zayavitel' i pravoobladatel' KFU (RU). – №2017663206; decl. 19.12.2017; publ. 16.02.2018, Reyestr programm dlya EVM.
4. Elizarov A.M., Lipachev E.K., and Khaydarov S.M. Automated System of Services for Processing of Large Collections of Scientific Documents // CEUR Workshop Proceedings. – 2016. – Vol. 1752. – P. 58–64.
5. Sylwestrzak W., Borbinha J., Bouche T., Nowiński A., and Sojka P. EuDML – Towards the European Digital Mathematics Library // In: Sojka P. (ed.): Towards a Digital Mathematics Library. Paris, France, July 7–8th, 2010. Masaryk University Press, Brno, Czech Republic, 2010. – P. 11–26. URL: <https://dml.cz/handle/10338.dmlcz/702569>.
6. Bouche T., Goutorbe C., Jorda J.-P., and Jost M. The EuDML Metadata Schema: Version 1.0 // In: Sojka P., Bouche T. (eds.): Towards a Digital Mathematics Library. Bertinoro, Italy, July 20–21st, 2011. Masaryk University Press, Brno, Czech Republic, 2011. P. 45–61. URL: <https://dml.cz/handle/10338.dmlcz/702602>.
7. Bouche T., and Rákosník J. Report on the EuDML External Cooperation Model // In: Kaiser K., Krantz S., Wegner B. (Eds.): Topics and Issues in Electronic Publishing, JMM, Special Session, San Diego, January 2013. – P. 99–108. URL: https://www.emis.de/proceedings/TIEP2013/07bouche_rakosnik.pdf.
8. Developing a 21st Century Global Library for Mathematics Research. Washington, D.C.: The National Academies Press, Washington, D.C (2014). arxiv.org/pdf/1404.1905, <https://www.nap.edu/catalog/18619/developing-a-21st-century-global-library-for-mathematics-research>
9. Elizarov A.M., Lipachev E.K., and Zuev D.S. Digital Mathematical Libraries: Overview of Implementations and Content Management Services // CEUR Workshop Proceedings. – 2017. – Vol. 2022. – P. 317–325.
10. Ion P.D.F., and Watt S.M. The Global Digital Mathematics Library and the International Mathematical Knowledge Trust // Geuvers H. et al. (Eds.): CICM

- 2017, LNAI 10383. 2017. – P. 56–69. https://doi.org/10.1007/978-3-319-62075-6_5.
11. Svidetel'stvo o gosudarstvennoy registratsii programmy dlya EVM. Programmnyy kompleks vydeleniya metadannykh iz kollektsey fiziko-matematicheskikh dokumentov, predstavlenykh v formate OpenXML / A.M. Elizarov, E.K. Lipachev, S.M. Khaydarov; zayavitel' i pravoobladatel' KFU (RU). – №2016616511; decl. 21.06.16; publ. 08.08.16, Reyestr programm dlya EVM.
 12. EuDML metadata schema specification (v2.0 – final). URL: <https://initiative.eudml.org/eudml-metadata-schema-specification-v20-final>.
 13. DBLP XML submission format guidelines. URL: <http://dblp.uni-trier.de/faq/dblpsubmission.xsd>.
 14. Elizarov A.M., and Lipachev E.K. Semanticheskiye metody i instrumenty elektronnoy matematicheskoy biblioteki Lobachevskii-DML // Nauchnyy servis v seti Internet: trudy XIX Vserossiyskoy nauchnoy konferentsii (18–23 september 2017, Novorossiysk). – Moscow: M.V. Keldysh Institute of Applied Mathematics, 2017. – P. 130–136. URL: <http://keldysh.ru/abrau/2017/73.pdf>. <https://doi.org/10.20948/abrau-2017-73>.
 15. Elizarov A.M., and Lipachev E.K. Lobachevskii DML: Towards a Semantic Digital Mathematical Library of Kazan University // CEUR Workshop Proceedings. – 2017. – Vol. 2022. – P. 326–333.