УДК 004.82+004.9

ФОРМИРОВАНИЕ МЕТАДАННЫХ ДЛЯ МЕЖДУНАРОДНЫХ БАЗ ЦИТИРОВАНИЯ В СИСТЕМЕ УПРАВЛЕНИЯ ЭЛЕКТРОННЫМИ НАУЧНЫМИ ЖУРНАЛАМИ

А.Н. Герасимов[1], А.М. Елизаров[2], Е.К. Липачёв[3]

Институт математики и механики им. Н.И. Лобачевского Казанского (Приволжского) федерального университета
[1]sav241@mail.ru, [2]amelizarov@gmail.com, [3]elipachev@gmail.com

Аннотация

Предложен алгоритм автоматического извлечения библиографических данных из однородного массива публикаций (в частности, выпусков научного журнала) и формирования блоков метаданных для экспорта в международные информационно-аналитические системы. Развиты методы интеграции платформы управления электронными научными журналами Open Journal Systems и международных баз научного цитирования.

Ключевые слова*: издательские системы, электронный научный журнал, интеграция электронных ресурсов, базы данных научного цитирования, экстракция метаданных*

ВВЕДЕНИЕ

В соответствии с современными международными стандартами публикация (в том числе, в электронной форме) научного журнала (текущего номера или выпуска с соответствующим набором статей) предполагает дальнейшую обработку публикуемых материалов информационно-аналитическими системами. К последним, например, относятся международные библиографические и реферативные базы данных Scopus (http://www.scopus.com/) и Web of Science (http://apps.webofknowledge.com/), а также российская национальная информационно-аналитическая система Российский индекс научного цитирования (РИНЦ, http://elibrary.ru/project_risc.asp). В течение последнего десятилетия они активно

используются для оценки научного уровня, как самих журналов, так и публикуемых ими статей с помощью целого набора показателей (импакт-факторы журналов, индексы цитирования по различным базам данных и др.) (см., например, [1–7]). В частности, индекс цитирования принят в научном мире как показатель значимости трудов ученого, представляет собой число ссылок на публикации этого ученого в реферируемых научных периодических изданиях и используется для оценки качества научной деятельности. Другой оценкой качества статей служит полнота библиографических ссылок. Как отмечено в [4], англоязычные статьи содержат в среднем около 30, а русскоязычные – около 10 ссылок. Такое количественное различие объясняется, в частности, сложившейся отечественной практикой издания преимущественно бумажных периодических изданий, предусматривающей ограничения на объем как самой публикации, так и количество пристатейных ссылок.

Блок метаданных любой научной публикации обязательно включает ее библиографическое описание (авторы, название, источник (например, журнал), год издания, том, номер, начальная и конечная страницы), авторское резюме (аннотация, реферат) и ключевые слова, а также различную дополнительную информацию. Таковой могут быть названия и места расположения организаций, от имени которых авторы представили свои материалы (аффилиация); фамилии ученых (членов редколлегий), представивших статью к публикации (как, например, в таких журналах, как Доклады академии наук и Lobachevskii Journal of Mathematics), указание текущей версии публикации (как в ставших сегодня модными «живых публикациях» [8, 9]) и другая информация. Отметим, что для размещения в международных информационно-аналитических системах неанглоязычных научных материалов необходимо дополнительно включать в блок метаданных список авторов, название работы, аннотацию, аффилиацию и ключевые слова на английском языке, а список литературы привести в транслитерации.

Для автоматизированного формирования блока метаданных сегодня с успехом могут быть использованы современные информационные технологии, особенно в тех случаях, когда они применяются для подготовки и издания самого научного журнала. В этом отношении передовые позиции занимают те издания, которые базируются на современных издательских системах (например, Open

Journal Systems – OJS, см., например, https://pkp.sfu.ca/ojs/, https://ru.wikipedia.org/wiki/Open_Journal_Systems) и специализированных программных пакетах, служащих для набора текстов статей [8]. К таким изданиям относятся, прежде всего, многие физико-математические и инженерно-технические журналы.

ОСОБЕННОСТИ ПОДГОТОВКИ БИБЛИОГРАФИЧЕСКОГО ОПИСАНИЯ НАУЧНЫХ ПУБЛИКАЦИЙ

Библиографическим описанием является совокупность библиографических сведений о документе, его составной части или группе документов, составленных по определенным правилам и необходимых для общей характеристики этого документа. Как правило, оформление библиографического описания публикаций выполняется в соответствии с существующими библиографическими правилами, установленными Государственными стандартами. Для России эти стандарты таковы (см., например, http://www.gosthelp.ru/gost/gost1560.html):

- ГОСТ 7.1-2003. «Библиографическая запись. Библиографическое описание»;
- ГОСТ 7.80-2001. «Библиографическая запись. Заголовок. Общие требования и правила составления»;
- ГОСТ 7.82-2001. «Библиографическая запись. Библиографическое описание электронных ресурсов»;
- ГОСТ 7.12-93. «Библиографическая запись. Сокращение слов на русском языке. Общие требования и правила».

Основные положения и правила стандартов основаны на принципах Международного стандарта библиографического описания (ISBD – International Standard Bibliographical Description) [10], принятые в системе международного библиографического учета мировых информационных документов и широко применяемые в национальных библиографических изданиях и библиотеках многих стран мира. Благодаря следованию единым стандартам облегчаются использование библиографи-

ческих материалов в мировой системе коммуникации, обмен информацией в машиночитаемой и электронной форме, преодолеваются языковые барьеры: записи, составленные в одной стране или на другом языке, могут быть легко поняты в другой стране или на другом языке.

Следование единым стандартам оформления значительно облегчает задачу обработки больших объемов разнородных электронных научных коллекций, что позволяет, как проводить аналитические исследования (выполнять запросы) по различным тематическим срезам, так и получать индексы цитирования [11]. Однако не все журналы в научном пространстве следуют единым стандартам, многие из них предлагают свои стилевые файлы, задающие общие правила оформления научной публикации и списка литературы, что затрудняет процесс автоматической обработки.

РЕГУЛЯРНЫЕ ВЫРАЖЕНИЯ В АЛГОРИТМАХ ИЗВЛЕЧЕНИЯ МЕТАДАННЫХ ИЗ КОЛЛЕКЦИЙ НАУЧНЫХ СТАТЕЙ

Как известно, большинство статей набирается с помощью офисных систем (например, MS Word, OpenOffice и другие) или систем, основанных на T_EX-нотации. Как правило, научные журналы принимают к рассмотрению и дальнейшей публикации только те статьи, которые подготовлены в соответствии с требованиями, четко сформулированными в каждом издании, в одном форматов .rtf, .doc, .docx, .odt и .tex. Выдвинутые требования регламентируют стилевые особенности оформления статей для журнала, а именно, порядок следования информационных блоков, оформление списка литературы и другие. Как правило, в статьях используются простые (с семантической точки зрения) средства структурирования. Это усложняет автоматическую обработку публикаций. Тем не менее, на основе анализа шрифтового выделения текста и порядка следования текстовых единиц (например, название, автор, аннотация и т. д.) и стандартных заголовков («Аннотация», «Ключевые слова», «Литература» и другие) можно определить структуру таких документов, что дает возможность проанализировать текст и автоматически

извлечь метаданные [12]. В рамках описанного подхода может быть предложен следующий алгоритм разбора библиографических списков:

– из текста статьи выделяется блок, содержащий список литературы: в качестве признака используются ключевые слова «References», «Список литературы», «Литература», «thebibliography» (для материалов в TEX-нотации);

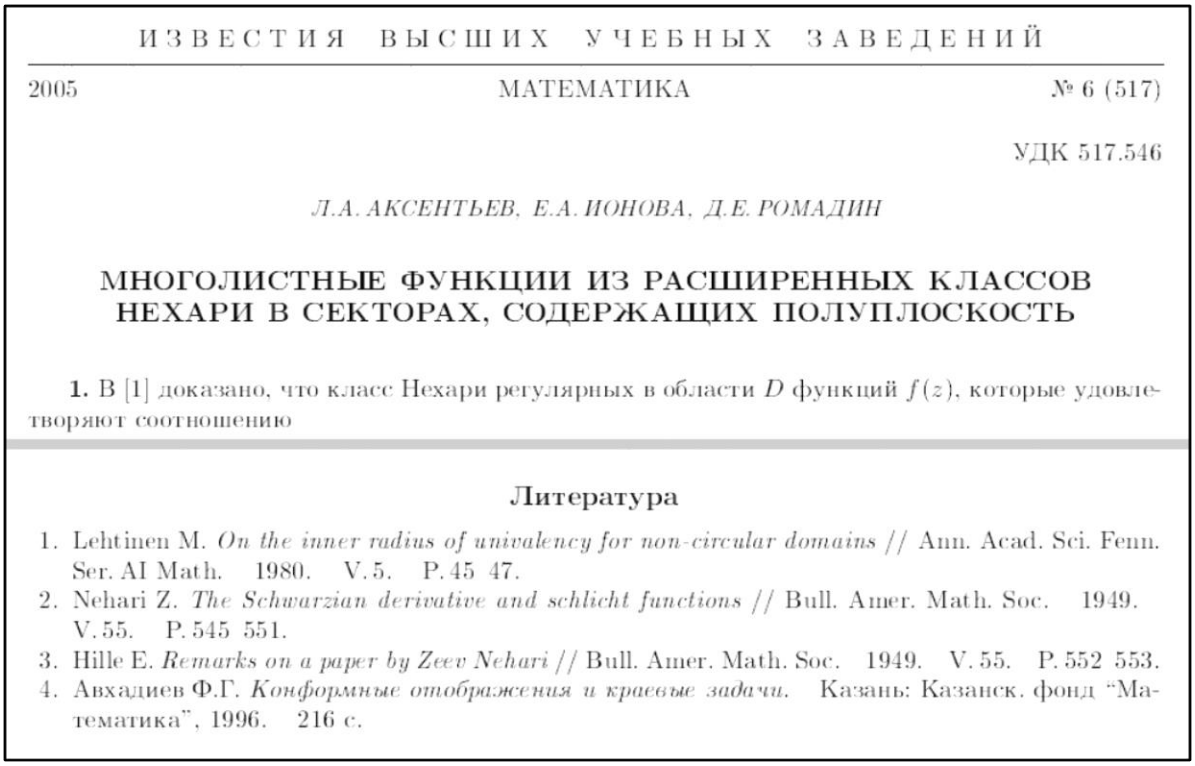
– извлекаются отдельные библиографические записи, признаками которых служат принятые в журнале правила оформления библиографических источников: квадратные скобки, нумерация, тег «bibitem» (для материалов в TEX-нотации);

– каждая библиографическая запись разделяется на элементы описания: список авторов, название и т. д.; для этого используется техника регулярных выражений (см., например, [14, 15]);

– для каждой коллекции подбирается система паттернов регулярных выражений, охватывающая возможные варианты написания библиографических источников;

– на заключительном этапе формируется XML-файл, содержащий метаописание коллекции статей, который используется для последующей конвертации в форматы баз научного цитирования.

Описанный алгоритм был апробирован на электронных коллекциях математических изданий: «Труды Математического центра им. Н.И. Лобачевского», «Ученые записки Казанского университета» и «Известия вузов. Математика» [16]. Общий объем коллекций – более 3 тыс. документов в форматах .pdf и .tex. Отметим, что эти документы имеют различную структуру, что потребовало разработки нескольких вариантов алгоритмов извлечения метаданных. Для работы с коллекцией документов каждого из указанных журналов подготовлена система паттернов, позволяющих выполнить извлечение и разбор библиографии.



ИЗВЕСТИЯ ВЫСШИХ УЧЕБНЫХ ЗАВЕДЕНИЙ

2005 МАТЕМАТИКА № 6 (517)

УДК 517.546

Л.А. АКСЕНТЬЕВ, Е.А. ИОНОВА, Д.Е. РОМАДИН

МНОГОЛИСТНЫЕ ФУНКЦИИ ИЗ РАСШИРЕННЫХ КЛАССОВ НЕХАРИ В СЕКТОРАХ, СОДЕРЖАЩИХ ПОЛУПЛОСКОСТЬ

1. В [1] доказано, что класс Нехари регулярных в области D функций $f(z)$, которые удовлетворяют соотношению

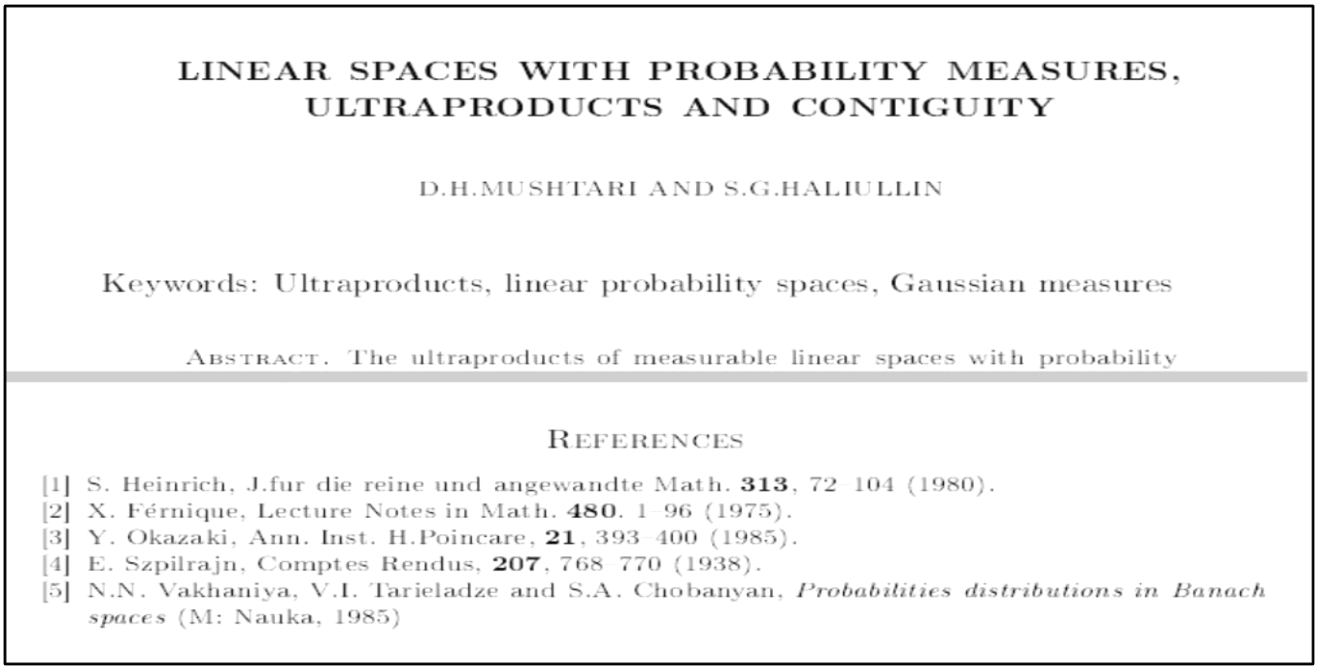
Литература

1. Lehtinen M. *On the inner radius of univalency for non-circular domains* // Ann. Acad. Sci. Fenn. Ser. AI Math. 1980. V. 5. P. 45 47.
2. Nehari Z. *The Schwarzian derivative and schlicht functions* // Bull. Amer. Math. Soc. 1949. V. 55. P. 545 551.
3. Hille E. *Remarks on a paper by Zeev Nehari* // Bull. Amer. Math. Soc. 1949. V. 55. P. 552 553.
4. Авхадиев Ф.Г. *Конформные отображения и краевые задачи.* Казань: Казанск. фонд "Математика", 1996. 216 с.

Рис. 1. Фрагмент статьи из журнала «Известия вузов. Математика», содержащий список авторов, название статьи и часть списка литературы

```
\setcounter{page}{3} \god{2005}
\nomer{6~(517)}
\udk{517.546}
\author{\it Л.А.\,Аксентьев, Е.А.\,Ионова, Д.Е.\,Ромадин}
\title{Многолистные функции из расширенных классов Нехари в секторах,
содержащих полуплоскость}
\date{}
\pagestyle{plain}
\begin{document}
\maketitle
\label{begin}
{\bf 1.} В [1] доказано, что класс Нехари регулярных в области $D$ функций
$f(z)$, которые удовлетворяют соотношению

\begin{thebibliography}{99}
\bibitem{1}
Lehtinen M. {\em On the inner radius of univalency for non-circular
domains} // Ann. Acad. Sci. Fenn. Ser.\,AI Math. -- 1980. -- V.\,5.
-- P.\,45--47.
\bibitem{2}
Nehari Z. {\em The Schwarzian derivative and schlicht functions} // Bull.
Amer. Math. Soc. -- 1949. -- V.\,55. -- P.\,545--551.
\bibitem{3}
Hille E. {\em Remarks on a paper by Zeev Nehari} // Bull. Amer. Math. Soc. --
1949. -- V.\,55. -- P.\,552--553.
\bibitem{4}
Авхадиев Ф.Г. {\em Конформные отображения и краевые задачи}. -- Казань:
Казанск. фонд ``Математика'', 1996. -- 216~с.
```

Рис. 2. Фрагмент TEX-кода статьи из журнала «Известия вузов. Математика»



LINEAR SPACES WITH PROBABILITY MEASURES, ULTRAPRODUCTS AND CONTIGUITY

D.H.MUSHTARI AND S.G.HALIULLIN

Keywords: Ultraproducts, linear probability spaces, Gaussian measures

Abstract. The ultraproducts of measurable linear spaces with probability

References

[1] S. Heinrich, J.fur die reine und angewandte Math. **313**, 72–104 (1980).
[2] X. Férnique, Lecture Notes in Math. **480**. 1–96 (1975).
[3] Y. Okazaki, Ann. Inst. H.Poincare, **21**, 393–400 (1985).
[4] E. Szpilrajn, Comptes Rendus, **207**, 768–770 (1938).
[5] N.N. Vakhaniya, V.I. Tarieladze and S.A. Chobanyan, *Probabilities distributions in Banach spaces* (M: Nauka, 1985)

Рис. 3. Фрагмент статьи из журнала «Lobachevskii Journal of Mathematics»

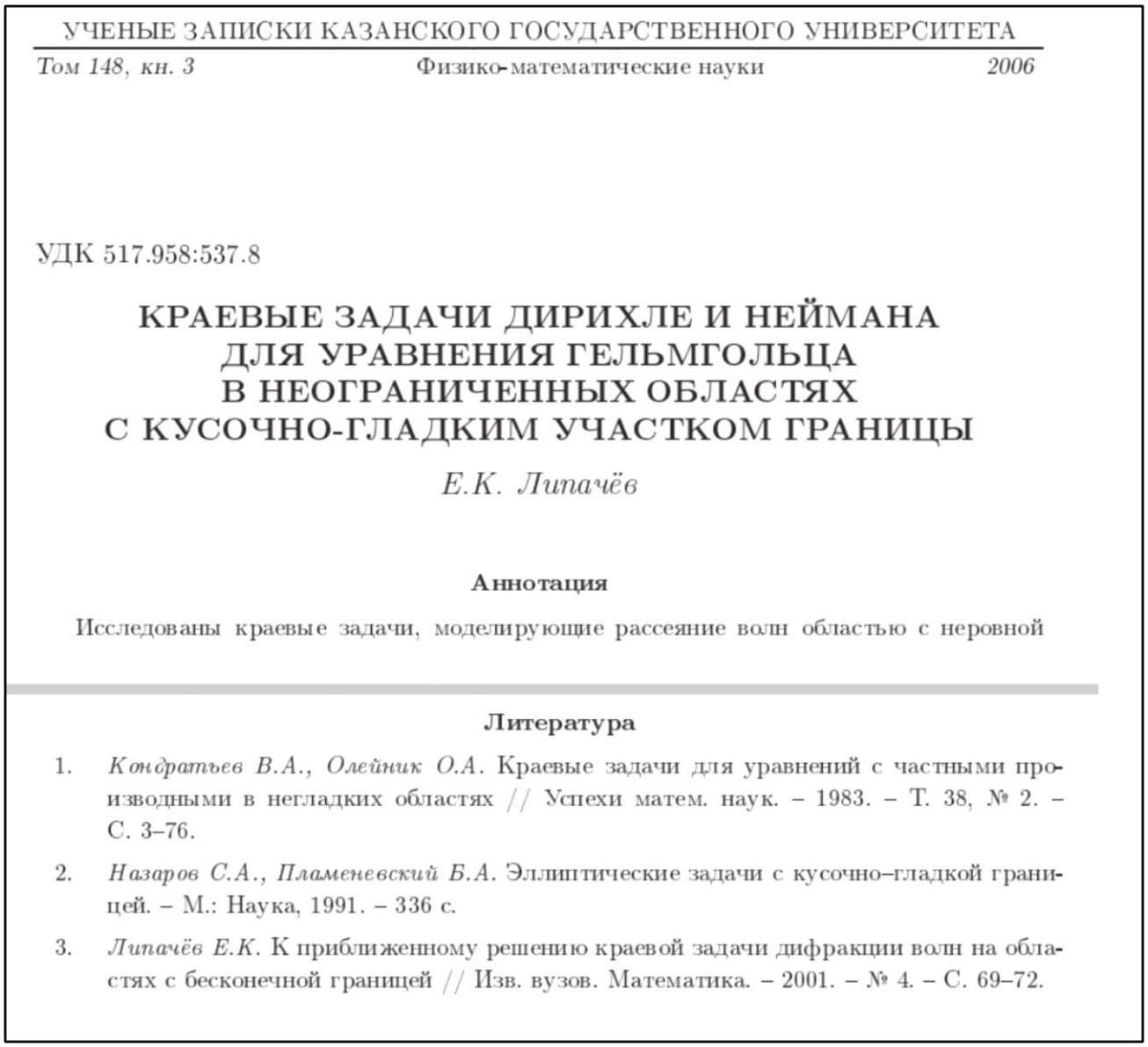
```
\begin{document}

\author{D.H.Mushtari and S.G.Haliullin}
\address{Kazan Federal University, 420008, Kremlevskaya, 18, Kazan, Russia}
\email{Samig.Haliullin@kpfu.ru}

\title{Linear spaces with probability measures, ultraproducts and contiguity}

\maketitle
\keywords{Keywords: Ultraproducts, linear probability spaces, Gaussian measures}


\begin{thebibliography}{99}
\bibitem{1} S. Heinrich, J.fur die reine und angewandte Math. \textbf{313}, 72
--104 (1980).
\bibitem{2} X. F\'ernique,  Lecture Notes in  Math. \textbf{480}. 1--96 (1975).
\bibitem{3} Y. Okazaki, Ann. Inst. H.Poincare, \textbf{21}, 393--400 (1985).
\bibitem{4} E. Szpilrajn, Comptes Rendus, \textbf{207}, 768--770 (1938).
\bibitem{5} N.N. Vakhaniya, V.I. Tarieladze and S.A. Chobanyan, \textit
{Probabilities distributions  in
Banach spaces} (M: Nauka, 1985)
```

Рис. 4. Запись приведенного ранее фрагмента статьи в TEX-нотации

Статьи журнала «Известия вузов. Математика» и «Lobachevskii Journal of Mathematics» размечены по правилам языка TEX с использованием стилевых инструкций, разработанных для данных журналов, в которых для библиографических списков используется окружение \thebibliography (см. рис. 1).

Для разбора библиографической записи использовалось регулярное выражение, где отличительным признаком было наличие окружения \bibitem для каждой записи:

@"item{.*?}(.*?){.em(.*?)}.*?[--|//](.*?)(\d\d\d\d).*?--(.*?)[bib|end";



УЧЕНЫЕ ЗАПИСКИ КАЗАНСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА

Том 148, кн. 3 Физико-математические науки *2006*

УДК 517.958:537.8

КРАЕВЫЕ ЗАДАЧИ ДИРИХЛЕ И НЕЙМАНА ДЛЯ УРАВНЕНИЯ ГЕЛЬМГОЛЬЦА В НЕОГРАНИЧЕННЫХ ОБЛАСТЯХ С КУСОЧНО-ГЛАДКИМ УЧАСТКОМ ГРАНИЦЫ

Е.К. Липачёв

Аннотация

Исследованы краевые задачи, моделирующие рассеяние волн областью с неровной

Литература

1. *Кондратьев В.А., Олейник О.А.* Краевые задачи для уравнений с частными производными в негладких областях // Успехи матем. наук. – 1983. – Т. 38, № 2. – С. 3–76.
2. *Назаров С.А., Пламеневский Б.А.* Эллиптические задачи с кусочно–гладкой границей. – М.: Наука, 1991. – 336 с.
3. *Липачёв Е.К.* К приближенному решению краевой задачи дифракции волн на областях с бесконечной границей // Изв. вузов. Математика. – 2001. – № 4. – С. 69–72.

Рис. 5. Фрагмент статьи из журнала «Ученые записки Казанского университета. Серия Физико-математические науки»

```
\UDK{517.958:537.8}

\ArticleNAME{КРАЕВЫЕ ЗАДАЧИ ДИРИХЛЕ И НЕЙМАНА \\
ДЛЯ УРАВНЕНИЯ ГЕЛЬМГОЛЬЦА В ОБЛАСТЯХ \\
С БЕСКОНЕЧНОЙ КУСОЧНО--ГЛАДКОЙ ГРАНИЦЕЙ }
\ArticleAUTHOR{Е.К.~Липачёв}
\ArticleHEAD{Краевые задачи Дирихле и Неймана}
\ArticleAUTHORHEAD{Е.К.~Липачёв}
\makeabstitle

\begin{abstract}
Исследованы краевые задачи, моделирующие рассеяние волн областью с
неровной границей.
```

```
\References{
\bibitem{kon}
{\it Кондратьев~В.А., Олейник~О.А.\/} Краевые задачи для уравнений
с частными производными в негладких областях
 // Успехи матем. наук. -- 1983. -- Т.~38, N~2. -- С.~3 -- 76.
\bibitem{nazplam} {\it Назаров~С.А., Пламеневский~Б.А.\/}
Эллиптические задачи с кусочно--гладкой границей. -- М.: Наука,
1991. -- 336~с.
\bibitem{lip2001} {\it Липачёв~Е.К.\/}
К приближенному решению краевой задачи дифракции волн на областях
с бесконечной границей // Изв. Вузов. Математика. -- 2001. -- N~4
(467). -- С.~69 -- 72.
```

Рис. 6. Запись приведенного ранее фрагмента статьи в T_EX-нотации

В журнале «Ученые записки Казанского университета. Серия Физико-математические науки» каждая запись списка выделялась окружением \bibitem, фамилии и имена авторов были выделены в окружение \it или \em, а весь блок библиографии – в \References. Было построено и применено регулярное выражение

@"item.*?{.[em|it](.*?)}(.*?)[--|//](.*?)(\d\d\d\d).*?--(.*?)[bib|end]"

В. И. Жегалов

Казанский (Приволжский) федеральный университет,
Институт математики и механики им. Н. И. Лобачевского,
vzhegalov@yandex.ru

ОБ ОДНОМ НАПРАВЛЕНИИ В ТЕОРИИ УРАВНЕНИЙ С ЧАСТНЫМИ ПРОИЗВОДНЫМИ

В течение ряда последних десятилетий ведется интенсивное исследование краевых задач, граничные условия в которых представляют собой соотношения, связывающие значения

ЛИТЕРАТУРА

1. Стеклов В. А. *Задача об охлаждении неоднородного твердого тела* // Сообщение Харьковского матем. об-ва, 1896. – Сер. 2. – Т. 5. – № 3–4. – С. 136–181.

2. Стеклов В. А. *Основные задачи математической физики.* – М.: Наука, 1983. – 432 с.

3. Франкль Ф. И. *Обтекание профилей газом с местной сверхзвуковой зоной, оканчивающейся прямым скачком уплотнения* // ПММ. – 1956. – Т. 20. – № 2. – С. 196–202.

Рис. 7. Фрагмент статьи из сборника «Труды Математического центра им. Н.И. Лобачевского»

```
\begin{document}
\ArticleAUTHOR{В.\,И.\,Жегалов \label{Zhegalov}}
\ArticleADRESS{Казанский (Приволжский) федеральный университет,
Институт математики и механики им. Н.И.~Лобачевского, \\
vzhegalov@yandex.ru}
\ArticleNAME{ОБ ОДНОМ НАПРАВЛЕНИИ В ТЕОРИИ УРАВНЕНИЙ С ЧАСТНЫМИ ПРОИЗВОДНЫМИ}
\makeabstitle
В течение ряда последних десятилетий ведется интенсивное исследование краевых задач,
граничные условия в которых представляют собой соотношения, связывающие значения
искомых функций в различных переменных точках, а также на линиях, расположенных
внутри рассматриваемой области.


\centerline{Л И Т Е Р А Т У Р А}
\smallskip
1.\;Стеклов\,В.\,А. {\it Задача об охлаждении неоднородного твердого тела}
// Сообщение Харьковского матем. об-ва, 1896.~-- Сер.~2.~-- Т.~5.~--
\No~3--4.~-- С.~136--181.

2.\;Стеклов\,В.\,А. {\it Основные задачи математической физики.}~--
М.: Наука, 1983.~-- 432~с.

3.\;Франкль\,Ф.\,И. {\it Обтекание профилей газом с местной сверхзвуковой зоной,
оканчивающейся прямым скачком уплотнения}
// ПММ.~-- 1956.~-- Т.~20.~-- \No~2.~-- С.~196--202.
```

Рис. 8. Запись приведенного ранее фрагмента статьи в T_EX-нотации

В сборнике «Труды Математического центра им. Н.И. Лобачевского» фамилии и имена авторов были выделены в окружение \it, а весь блок библиографии – по ключевому слова «Литература». Каждая запись выделялась с помощью регулярного выражения

@"[0-9](.*?){\\it(.*?)}(.*?)(\d{4}).*?([CP].~\d+--\d+|\d+~[cp])."

Russian Digital Libraries
JOURNAL

ISSN 1562-5419

2014 | Volume 17| Issue 2

A Model for Integrating the Publication and Preservation of Journal Articles

Kevin S. Hawkins

Abstract

There are policy, technical, and workflow gaps in library efforts to preserve online journal literature. Since libraries are increasingly involved in journal publishing, HathiTrust, a shared preservation-quality digital repository, is a natural place to archive and provide access to journal literature to ensure its long-term preservation and discoverability. The University of Michigan Library is funding the creation of mPach, an open-source, end-to-end publishing system in which archiving in HathiTrust happens as a byproduct of publication rather than being carried out after the fact. The architecture of mPach, its envisioned workflow, and plans for creating a shared infrastructure for publishing open-access journals are all summarized.

Key words: magazines published online, digital repository HathiTrust, open-access journals, the system mPach of the complete a publication cycle.

References

[1] Sadie L. Honey. Preservation of electronic scholarly publishing: an analysis of three approaches. Portal: libraries and the academy, 5(1):59-75, Jan. 2005.

[2] NISO SERU Standing Committee. SERU: a shared electronic resource understanding: a recommended practice of the National Information Standards Organization. Baltimore: National Information Standards Organization. 2012. http://www.niso.org/publications/rp/RP-7-2012_SERU.pdf.

[3] Lots of Copies Keeps Stuff Safe. http://www.lockss.org/.

[4] CLOCKSS. http://www.clockss.org/.

[5] Portico. http://www.portico.org/.

[6] National Library of the Netherlands and Elsevier Science make digital preservation history: permanent digital archive assures perpetual accessibility of scientific heritage. August 20, 2002. http://www.kb.nl/en/news/news-archive-2002/national-library-of-the-netherlands-and-elsevier-science-make-digital-preservation-history.

Рис. 9. Страница журнала «Электронные библиотеки»

На сегодняшний день для загрузки статей в базу данных Российского индекса научного цитирования (РИНЦ) используется онлайн-программа разметки Articulus (http://elibrary.ru/projects/contracts/publisher/messages/messages.asp), предназначенная для подготовки выпусков журналов и книг в формате XML [17, 18]. Используя систему Articulus, разметчик вынужден с помощью ручного ввода заполнять поля метаданными, что требует от него значительных ресурсов (человеко-часов). При подготовке проведения крупных конференций с большим количеством материалов и сжатыми сроками размещения на сайтах соответствующих

электронных ресурсов, когда о ручной обработке не может идти речи, задача становится особенно актуальной. Естественно, что традиционными методами оперативно выполнить эту работу невозможно.

АЛГОРИТМ ФОРМИРОВАНИЯ БЛОКА МЕТАДАННЫХ

Автоматизация процесса формирования блока метаданных выполнялась в несколько этапов на массивах естественно-научных журналов, перечисленных выше, и гуманитарного журнала «Вестник КазГУКИ». Статьи предварительно проверялись на соответствие стилевым требованиям журналов, далее проводились обработка библиографических данных и их проверка на соответствие единому стандарту.

УДК 001+002+004

М.С.Галявиева, А.М.Елизаров, Р.У.Елизарова, Е.К.Липачёв

ГУМАНИТАРНЫЕ НАУЧНЫЕ ЖУРНАЛЫ:МЕТОДЫ И ИНСТРУМЕНТЫ ИНТЕГРАЦИИ В НАУЧНОЕ ИНФОРМАЦИОННОЕ ПРОСТРАНСТВО [1]

Обсужден подход к интеграции в научное информационное пространство гуманитарных научных журналов, базирующийся на использовании современных программных платформ управления электронными публикациями и электронными изданиями.

Рассмотрены основные характеристики Open Journal Systems (OJS) – одной из самых распространенных информационных систем автоматизации полного цикла подготовки и издания электронных научных журналов. Указаны отличительные особенности OJS и основные ее преимущества. Сделаны выводы о возможностях использования системы OJS для издания гуманитарного научного журнала по любой тематике и из любой предметной области, а также дополнения функционала системы (при необходимости) на основании плагинов, учитывающих специфику издаваемого журнала.

Охарактеризовано множество российских научных журналов, использующих OJS. Описана платформа управления научными журналами science.tatarstan.ru, развернутая в Республике Татарстан и базирующаяся на OJS. Представлен подход Article-level metrics, используемый для оценки научной и социальной значимости отдельно взятой научной статьи.

Ключевые слова: Веб, интернет, гуманитарные научные журналы, научное информационное пространство, информационные системы управления электронными научными журналами, Open Journal Systems, платформа science.tatarstan.ru, информетрия, альтметрия, Article-level metrics

Рис. 10. Страница из журнала «Вестник КазГУКИ» содержит основные метаданные

Для извлечения метаданных автоматически производился разбор файла статьи: файл преобразовывался в формат .docx, который основан на стандарте [19], построенном на базе стандартов XML, ZIP и XML-Schema. С использованием структуры .docx-файла извлекался основной компонент документа – файл *document.xml*. На рис. 11 приведен фрагмент PHP-скрипта, предназначенного для выделения библиографического списка из текста статьи.

```
//извлекаем document.xml из каждого docx файла в коллекции.
foreach($files as $file){
    $zip = new ZipArchive;
    $res = $zip->open($file);
    $zip->extractTo('./tmp', array('word/document.xml'));
    $zip->close(); $file=iconv("cp1251","UTF-8", $file);
    //echo $file.'<br/>';
    $file=str_replace("docx", "pdf",$file);
    $xml = new DOMDocument; $xml->load('tmp/word/document.xml')
    $w_ps=$xml->getElementsByTagNameNS(
    'http://schemas.openxmlformats.org/wordprocessingml/2006/main','p');
    //цикл по абзацам документа
    while($k<$w_ps->length){
        ...
        //Поиск библиографии
        if(preg_match("[0-9](.*?)//(.*?)(\d{4})(.*?)(P.|C.)(.*?)(\.)" ,
        $w_ps->item($k)->nodeValue )){
    }
...
```

Рис. 11. Фрагмент скрипта выделения библиографического списка

Отличительными признаками блока библиографии журнала является наличие библиографических записей на русском и английском языках, заголовка «Литература» и следование инициалов автором после фамилии. В силу перечисленных особенностей блок библиографии выделялся с помощью регулярного выражения, учитывающего особенности оформления (см. рис. 12).

Литература

1. D'Arcy P. CIO strategies for consumerization: the future of enterprise mobility. – Dell Power Solutions Special Issue, Dell Inc., 2012. – P. 22-25. – URL: http://www.dell.com/Learn/us/en/555/power-solutions-magazine-2012-special-issue.

2. Gantz J., Reinsel D. The Digital Universe in 2020: big data, bigger digital shadows, and biggest growth in the Far East. – IDC Digital Universe Study, December 2012. – URL: http://www.emc.com/leadership/digitaluniverse/iview/index.htm.

3. Елизаров А.М., Зуев Д.С., Липачёв Е.К. Свободно распространяемые системы управления электронными научными журналами и технологии электронных библиотек// Тр. 15-й Всерос. науч. конф. «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» – RCDL-2013, Ярославль: ЯрГУ, 2013. – С. 227-236.

4. Елизаров А.М., Липачёв Е.К. Системы интеграции электронной научно-образовательной информации и повышение поисковой

1. D'Arcy P. CIO strategies for consumerization: the future of enterprise mobility. – Dell Power Solutions Special Issue, Dell Inc., 2012. – P. 22-25. – URL: http://www.dell.com/Learn/us/en/555/power-solutions-magazine-2012-special-issue.

2. Gantz J., Reinsel D. The Digital Universe in 2020: big data, bigger digital shadows, and biggest growth in the Far East. – IDC Digital Universe Study, December 2012. – URL: http://www.emc.com/leadership/digitaluniverse/iview/index.htm.

3. Elizarov A.M., Zuev D.S., Lipachjov E.K. Svobodno rasprostranjaemye sistemy upravlenija jelektronnymi nauchnymi zhurnalami i tehnologii jelektronnyh bibliotek// Tr. 15-j Vseros. nauch. konf. «Jelektronnye biblioteki: perspektivnye metody i tehnologii, jelektronnye kollekcii» – RCDL-2013, Jaroslavl': JarGU, 2013. – S. 227-236.

4. Elizarov A.M., Lipachjov E.K. Sistemy integracii jelektronnoj nauchno-obrazovatel'noj informacii i

Рис. 12. Оформление списка литературы в журнале «Вестник КазГУКИ»

Для автоматического разбора библиографии было использовано несколько шаблонов регулярных выражений. Отличительным признаком статьи считалось наличие знаков "//" в библиографическом описании. В этом случае применялось регулярное выражение, в котором отдельные блоки метаданных выделялись в группы:

@"[0-9](.*?)//(.*?)(\d{4})(.*?)(P.|C.)(.*?)(\.)

Приведенное регулярное выражение позволило выделить такие группы, как список авторов, номер тома, страницы и другие блоки библиографического описания. Далее проводился разбор каждой группы, в частности, осуществлялось выделение отдельного автора из списка авторов статьи.

Для разбора библиографического описания книг использовался шаблон

@"([0-9]\.)(.*?)(\.)(.*?)(\.)(.*?)(\d{4})(.*?)(c.|p.)

В качестве следующего этапа улучшения алгоритма рассматривается возможность программного построения регулярных выражений на основе шаблона представленных материалов. В случае неудовлетворительного результата и указания ошибочных блоков предусмотрена возможность повторного построения выражений.

ИМПОРТ МЕТАДАННЫХ В БАЗЫ ДАННЫХ НАУЧНОГО ЦИТИРОВАНИЯ

Необходимым условием загрузки метаданных в международные информационно-аналитические базы является оформление пристатейной библиографии на латинице [5]. Соответствующий алгоритм реализован как веб-приложение, позволяющее загружать файлы статей в форматах .tex, .doc и в результате получить XML-файл, содержащий метаданные, структурированные согласно заданному формату выбранной базы цитирования. Приложение позволяет выбрать группу файлов указанием пути к папке с документами и автоматизировать процесс выделения метаданных статей целых сборников или коллекций документов, имеющих схожее форматирование, например, выпусков журнала или статей конференций. С помощью разработанного веб-приложения, реализованного на языке PHP, были получены XML-файлы, содержащие блоки метаданных, записанных в соответствии с правилами РИНЦ.

На рис. 13 приведен фрагмент XML-файла, содержащий метаданные статьи (см. рис. 10):

```
<article>
    <pages>86-96</pages>
    <artType>PRC</artType>
    <authors>
        …
        <author num="002" id="2521">
            <individInfo lang="RUS">
            <surname>Елизаров</surname>
            <initials>А.М.</initials>
            <orgName>Казанский (Приволжский) федеральный университет</orgName>
            </individInfo>
        </author>
        …
        <author num="004" id="10719">
            <individInfo lang="RUS">
            <surname>Липачев</surname>
            <initials>Е.К.</initials>
            <orgName>Казанский (Приволжский) федеральный университет</orgName>
            </individInfo>
        </author>
    </authors>
<artTitles>
    <artTitle lang="RUS">ГУМАНИТАРНЫЕ НАУЧНЫЕ ЖУРНАЛЫ:МЕТОДЫ И ИНСТРУМЕНТЫ ИНТЕГРАЦИИ В НАУЧНОЕ ИНФОРМАЦИОННОЕ ПРОСТРАНСТВО
    </artTitle>
    <artTitle lang="ENG">HUMANITARIAN SCIENTIFIC JOURNALS: METHODS AND TOOLS OF INTEGRATION IN SCIENTIFIC INFORMATION SPACE
    </artTitle>
</artTitles>
```

Рис. 13. XML-файл метаданных в формате РИНЦ

В документе, приведенном выше, атрибут id тега author есть уникальный регистрационный номер автора, присвоенный при регистрации. Заполнение данного атрибута является обязательным для автоматической привязки статьи к ее автору. Для получения уникального номера, а также дополнительной информации (представление на английском языке) был реализован отдельный модуль поиска в системе РИНЦ, который имитирует работу пользователя в веб-браузере.

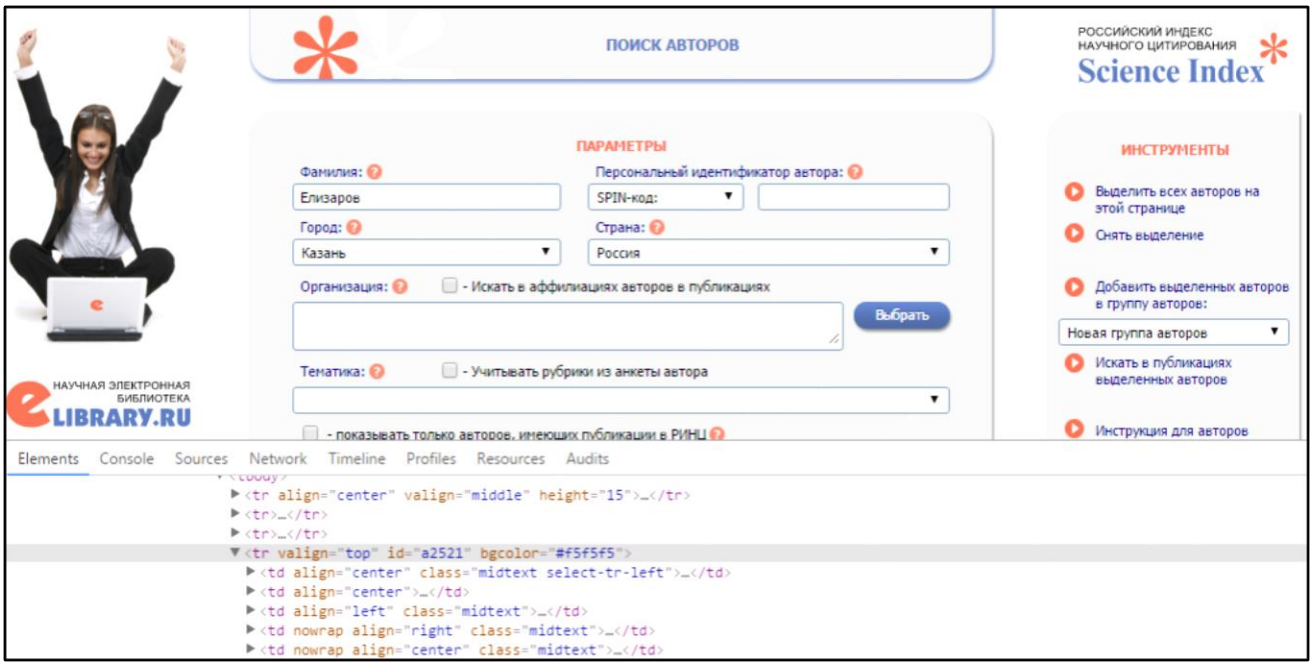


Рис. 14. Поиск автора в системе РИНЦ.

Следующий фрагмент (рис. 15) содержит часть списка литературы статьи (см. рис. 12), размеченную в автоматическом режиме:

```
<references>
    <reference>D'Arcy P. CIO strategies for consumerization: the future of enter-prise
    mobility. – Dell Power Solutions Special Issue, Dell Inc., 2012. – P. 22-25. – URL:
    http://www.dell.com/Learn/us/en/555/powersolutions-magazine-2012-special-issue.
    </reference>
    ...
    <reference>Ахметов Д.Ю., Елизаров А.М., Липачев Е.К. Система автома-тизации редакционных
    процессов на платформе электронных научных жур-налов // Ученые записки ИСГЗ. – 2014. –
    № 1-2 (12). – С. 228-233. </reference>
    ...
</references>
```

Рис. 15. Результат автоматической разметки списка литературы

Фрагмент XML-документа, построенного с разбиением списка литературы (см. рис. 12) на библиографические элементы, представлен на рис. 16.

```
<references>
    <reference id=1>
        <authors>
            <author>D'Arcy P.</author>
        </authors>
        <title-paper> CIO strategies for consumerization: the future of enterprise mobility.</title-paper>
        <year>2012</year>
        <pages>22-25</pages>
        <inf-bib>CIO strategies for consumerization: the future of enterprise mobility. - Dell
        Power Solutions Special Issue, Dell Inc., 2012. - P. 22-25. - URL:
        http://www.dell.com/Learn/us/en/555/powersolutions-magazine-2012-special-issue.</inf-bib>
    </reference>
    ...
    <reference id=20>
        <authors>
            <author>Ахметов Д.Ю.</author>
            <author>Елизаров А.М.</author>
            <author>Липачёв Е.К.</author>
        </authors>
        <title-paper>Свободно распространяемые системы управления электронными научными журналами
        и технологии электронных библиотек</title-paper>
        <year>2014</year>
        <pages>228-233</pages>
        <inf-bib>Система автоматизации редакционных процессов на платформе электронных научных
        журналов // Ученые записки ИСГЗ. - 2014. - № 1-2 (12). - С. 228-233.</inf-bib>
    </reference>
</references>
```

Рис. 16. Результат автоматической разметки списка литературы с разбиением на библиографические элементы

СЕРВИСЫ ИНТЕГРАЦИИ ИНФОРМАЦИОННЫХ СИСТЕМ УПРАВЛЕНИЯ ЭЛЕКТРОННЫМИ ЖУРНАЛАМИ И БАЗ НАУЧНОГО ЦИТИРОВАННИЯ

Все больше журналов начинают использовать различные системы управления электронными научными журналами, такие, как Open Journal System (OJS), ePublishing Toolkit, Ambra Publishing System и другие. Наибольшую популярность получила система OJS, являющаяся онлайн издательской системой полного цикла, что связано с наличием достаточного функционала и доступностью к его расширению [18, 12].

Внедрение платформ управления бизнес-процессами научного журнала позволяет автоматизировать наиболее трудоемкие рабочие процессы, в том числе дает возможность интеграции журнала с различными информационными системами. Платформа OJS обладает модулями экспорта метаданных в форматах:

- XML, по шаблону native.dtd;
- Erudit, определенном в виде DTD;
- CrossRef XML;
- PubMed XML для индексирования MEDLINE;

- XML для архивации в DOAJ.

В качестве расширения базовой функциональности платформы OJS была выполнена реализация алгоритмов экстракции метаданных в виде отдельного модуля, что позволило выгружать XML-файлы, сформированные в соответствии с заранее заданным форматами различных внешних информационных систем. Данный модуль представлен в виде отдельного веб-сервиса, записанного на языке WSDL (Web Services Description Language) [21] и предоставляющго доступ к журналам по протоколу SOAP (Simple Object Access Protocol) [22]. Для поддержки мультиязычности при выгрузке метаданных был реализован отдельный модуль, позволяющий выполнять транслитерацию в автоматическом режиме.

Предложенный алгоритм активно внедряется в виде веб-сервиса в информационную систему управления научными публикациями Science.Tatarstan.ru [21], при этом используются программные инструменты интеграции с международными базами цитирования системы Open Journal System [24, 25].

ЗАКЛЮЧЕНИЕ

Метаданные являются важным элементом научных публикаций, а их наличие и полнота очень важны для ориентации в научном пространстве и учета в аналитических системах. Подход, представленный выше, является шагом в развитии взаимодействия информационных систем научного характера, а внедрение платформ управления научными журналами позволит автоматизировать процесс загрузки метаданных научного контента в базы научного цитирования.

Благодарности

Работа выполнена при финансовой поддержке Российского гуманитарного научного фонда, проект 14-03-12004, а также Российского фонда фундаментальных исследований и Правительства Республики Татарстан в рамках научного проекта № 15-47-02472.

СПИСОК ЛИТЕРАТУРЫ

1. *Кириллова О.В.* Подготовка научных журналов по международным стандартам и требованиям индексов цитирования. URL: http://lib.ss-au.ru/uploaded/Publ/IC%20and%20science%20journals.pdf.

2. *Кириллова О.В.* О системе включения журналов в БД Scopus: основные требования и порядок представления. URL: http://www.webci-tation.org/ 68vOlqztg.

3. *Кириллова О.В.* Редакционная подготовка журналов по международным стандартами требованиям глобальных индексов цитирования. URL: http://www.mgsu.ru/science/Nssled_i_innovac_deyat/UNP/naukometriya/redakcion-naya_podgotovka_zhurnalov.pdf.

4. *Гусев А.Л.* Индексы цитирования и аналитический аппарат современной издательской платформы // Альтернативная энергетика и экология. 2013. №4. С. 87-91.

5. *Кириллова О.В.* Критерии отбора и рекомендации по подготовке журнала в индекс цитирования Scopus. URL: http://fano.gov.ru/common/upload/ library/2014/12/main/kriterii_journ-als.pdf.

6. *Когаловский М.Р.* Онтологическое аннотирование библиографических ссылок в научных публикациях и его использование в наукометрии// Информационные ресурсы России. 2013. № 5. С. 5-13.

7. *Когаловский М.Р., Паринов С.И.* Новый источник данных для наукометрических исследований// Труды 15-й Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» – RCDL-2013, Ярославль, Россия, 14–17 октября 2013 г. С. 107-117.

8. *Паринов С.И., Когаловский М.Р.* Технология поддержки электронных научных публикаций как «живых» документов // Труды XI Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции». — Петрозаводск: Карельский научный центр РАН, 17–21 сентября 2009. С. 53-58.

9. *Горбунов-Посадов М.М.* Живая публикация // Открытые системы. СУБД. 2011. № 4. С. 48-49.

10. Оформление библиографического списка. URL: http://lib.samgtu.ru/ making_list.

11. *Афонин С.А., Бахтин А.В., Бухонов В.Ю., Васенин В.А., Ганкин Г.М., Гаспарянц А.Э., Голомазов Д.Д., Иткес А.А., Козицын А.С., Тумайкин И.Н., Шапченко К.А.*

Интеллектуальная система тематического исследования научно-технической информации (ИСТИНА). Под ред. академика В.А. Садовничего. М.: Издательство Московского университета, 2014. 262 с.

12. *Елизаров А.М., Зуев Д.С., Липачёв Е.К.* Информационные системы управления электронными научными журналами // Научно-техническая информация. Серия 1: Организация и методика информационной работы. 2014. № 3. С. 31-38.

13. *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.

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.

15. *Гойвертс Я., Левитан С.* Регулярные выражения. Сборник рецептов. СПБ: Символ-Плюс, 2013. 608 с.

16. *Герасимов А.Н., Елизаров А.М., Липачёв Е.К.* Паттерны регулярных выражений в алгоритмах семантической обработки коллекций математических документов // Труды Казанской школы по компьютерной и когнитивной лингвистике. Казань, 2014. Вып. 16. С. 43-45.

17. *Глухов В.А.* Книжная коллекция. О размещении книг, сборников и материалов конференций в Российском индексе научного цитирования: Научная электронная библиотека. URL: http://www.library.spbu.ru/blog/wp-content/uploads/2014/12/books_Glukhov.pdf.

18. *Мбого И.А., Прокудин Д.Е., Чугунов А.В.* Комплексная интеграция цифровых коллекций в информационное пространство научных исследований // Технологии информационного общества в науке, образовании и культуре: сборник научных статей. Материалы XVII Всероссийской объединенной конференции «Интернет и современное общество» IMS-2014, Санкт-Петербург, 19 – 20 ноября 2014 г. С. 48-53.

19. Standard ECMA-376: Office Open XML File Formats. URL: http://www.ecmainternational.org/publications/standards/Ecma-376.htm.

20. *Willinsky J., Stranack K., Smecher A., MacGregor J.* Open Journal Systems: a complete guide to online publishing. Simon Fraser University Library, 2010. URL: https://pkp.sfu.ca/ojs/docs/userguide/2.3.3/index.html.

21. Web Services Description Language (WSDL) Version 2.0. W3C Recommendation. – http://www.w3.org/TR/2007/REC-wsdl20-20070626.

22. *Snell J., Tidwell D., Kulchenko P.* Programming Web Services with SOAP. O'Reilly Media, 2001. 262 p.

23. *Ахметов Д.Ю., Грачев А.О., Герасимов А.Н., Елизаров А.М., Липачёв Е.К.* Облачная платформа поддержки электронных научных изданий // Учёные записки Института социальных и гуманитарных знаний. 2014. № 1 (12), ч. 1. С. 13-19.

24. *Stranack K.* Getting found, staying found, increasing impact. Enhancing readership and preserving content for OJS journals // Public Knowledge Project. 2006. 40 p.

25. Importing and Exporting Data with OJS // URL: http://www.informatica.si/manual/ojs-import-export.pdf.

SUBSYSTEM OF FORMATION METADATA FOR SCIENCE INDEX DATABASES ON MANAGEMENT PLATFORM ELECTRONIC SCIENTIFIC JOURNALS

A.N. Gerasimov[1], A.M. Elizarov[2], E.K. Lipachev[3]

N.I. Lobachevskii Institute of Mathematics and Mechanics. Kazan Federal University

[1]sav241@mail.ru, [2]amelizarov@gmail.com, [3]elipachev@gmail.com

Abstract

An algorithm for automatic extraction of bibliographic data from a one-dimensional array of publications (in particular, the issues of the scientific journal) and the formation of metadata blocks for export to international information-analytical system. The methods of integration management platform electronic scientific journals (Open Journal Systems) and Science Citation Index.

***Keywords**: publishing systems, digital scientific journal, the integration of electronic resources, databases, scientific citation, metadata extraction*

REFERENCES

1. *Kirillova O.V.* Podgotovka nauchnykh zhurnalov po mezhdunarodnym standartam i trebovaniyam indeksov tsitirovaniya. URL: http://lib.ss-au.ru/uploaded/Publ/IC%20and%20science%20journals.pdf.

2. *Kirillova O.V.* O sisteme vklyucheniya zhurnalov v BD Scopus: osnovnye trebovaniya i poryadok predstavleniya. URL: http://www.webci-tation.org/68vOlqztg.

3. *Kirillova O.V.* Redaktsionnaya podgotovka zhurnalov po mezhdunarodnym standartami trebovaniyam global'nykh indeksov tsitirovaniya. URL: http://www.mgsu.ru/science/Nssled_i_innovac_deyat/UNP/naukometriya/redakcionnaya_podgotovka_zhurnalov.pdf.

4. *Gusev A.L.* Indeksy tsitirovaniya i analiticheskij apparat sovremennoj izdatel'skoj platformy // Al'ternativnaya ehnergetika i ehkologiya. 2013. №4. S. 87-91.

5. *Kirillova O.V.* Kriterii otbora i rekomendatsii po podgotovke zhurnala v indeks tsitirovaniya Scopus. URL: http://fano.gov.ru/common/upload/library/2014/12/ main/kriterii_journ-als.pdf.

6. *Kogalovskij M.R.* Ontologicheskoe annotirovanie bibliograficheskih ssylok v nauchnyh publikacijah i ego ispol'zovanie v naukometrii// Informacionnye Resursy Rossii. 2013. № 5. S. 5-13.

7. *Kogalovskij M.R., Parinov S.I.* Novyj istochnik dannyh dlja nauko-metricheskih issledovanij// Trudy 15-j Vserossijskoj nauchnoj konferencii «Jelektronnye biblioteki: perspektivnye metody i tehnologii, jelektronnye kollekcii» — RCDL-2013, Jaroslavl', Rossija, 14–17 oktjabrja 2013 g. S. 107-117.

8. *Parinov S.I., Kogalovskij M.R.* Tekhnologiya podderzhki ehlektronnykh nauchnykh publikatsij kak «zhivykh» dokumentov // Trudy XI Vserossij-skoj nauchnoj konferentsii «EHlektronnye biblioteki: perspektivnye meto-dy i tekhnologii, ehlektronnye kollektsii». Petrozavodsk: Karel'skij nauchnyj tsentr RAN, 17–21 sentyabrya 2009. S. 53-58.

9. *Gorbunov-Posadov M.M.* Zhivaya publikatsiya // Otkrytye sistemy. SUBD. 2011. № 4. S. 48-049.

10. Oformlenie bibliograficheskogo spiska. URL: http://lib.samgtu.ru/making_list.

11. *Afonin S.A., Bahtin A.V., Buhonov V.Ju., Vasenin V.A., Gankin G.M., Gasparjanc A.Je. , Golomazov D.D., Itkes A.A., Kozicyn A.S., Tumajkin I.N., Shapchenko K.A.* Intellektual'naja sistema tematicheskogo issledovanija nauchno-tehnicheskoj informacii (ISTINA). Pod red. akademika V.A. Sadovnichego. M.: Izdatel'stvo Moskovskogo universiteta, 2014. 262 s.

12. *Elizarov A.M., Zuev D.S., Lipachev E.K.* Informatsionnye sistemy upravleniya ehlektronnymi nauchnymi zhurnalami // Nauchno-tekhnicheskaya informatsiya. Seriya 1: Organizatsiya i metodika informatsionnoj raboty. 2014. № 3. S. 31-38.

13. *Elizarov A.M., Lipachev E.K., Khokhlov Yu.E.* Semanticheskie metody strukturirovaniya matematicheskogo kontenta, obespechivayushhie rasshirennuyu poiskovuyu funktsional'nost' // Informatsionnoe obshhestvo. 2013. № 1-2. S. 83-92.

14. *Elizarov A.M., Lipachev E.K., Nevzorova O.A., Solov'ev V.D.* Metody i sredstva semanticheskogo strukturirovaniya ehlektronnykh matematicheskikh dokumentov // Doklady Akademii nauk. 2014. T. 457. № 6. S. 642-645.

15. *Gojverts J., Levithan C.* Reguljarnye vyrazhenija. Sbornik receptov. SPB: Simvol-Pljus, 2013. 608 p.

16. *Gerasimov A.N., Elizarov A.M., Lipachev E.K.* Patterny regulyarnykh vyrazhenij v algoritmakh semanticheskoj obrabotki kollektsij matematicheskikh dokumentov // Trudy Kazanskoj shkoly po komp'yuternoj i kogni-tivnoj lingvistike. Kazan', 2014, Vyp. 16. S. 43-45.

17. *Gluhov V.A.* Knizhnaja kollekcija. O razmeshhenii knig, sbornikov i materialov konferencij v Rossijskom indekse nauchnogo citirovanija: Nauchnaja jelektronnaja biblioteka. URL: http://www.library.spbu.ru/blog/wp-content/uploads/2014/12/books_Glukhov.pdf.

18. *Mbogo I.A., Prokudin D.E., Chugunov A.V.* Kompleksnaya integratsiya tsifrovykh kollektsij v informatsionnoe prostranstvo nauchnykh issledovanij // Tekhnologii informatsionnogo obshhestva v nauke, obrazovanii i kul'ture: sbornik nauchnykh statej. Materialy XVII Vserossijskoj ob"edinennoj konferentsii «Internet i sovremennoe obshhestvo» IMS-2014, Sankt-Peterburg, 19 – 20 noyabrya 2014 g. S. 48-53.

19. Standard ECMA-376: Office Open XML File Formats. URL: http://www.ecmainternational.org/publications/standards/Ecma-376.htm.

20. *Willinsky J., Stranack K., Smecher A., MacGregor J.* Open Journal Systems: A complete guide to online publishing. Simon Fraser University Library, 2010. URL: https://pkp.sfu.ca/ojs/docs/userguide/2.3.3/index.html.

21. Web Services Description Language (WSDL) Version 2.0. W3C Recommendation. URL: http://www.w3.org/TR/2007/REC-wsdl20-20070626.

22. *Snell J., Tidwell D., Kulchenko P.* Programming Web Services with SOAP. O'Reilly Media, 2001. 262 p.

23. *Akhmetov D.Yu., Grachev A.O., Gerasimov A.N., Elizarov A.M., Lipachev E.K.* Oblachnaya platforma podderzhki ehlektronnykh nauchnykh izdanij // Uchyonye zapiski Instituta sotsial'nykh i gumanitarnykh znanij. 2014. № 1 (12). ch. 1. S. 13-19.

24. *Stranack K.* Getting found, staying found, increasing impact. Enhancing Readership and Preserving Content for OJS Journals // Public Knowledge Project. 2006. 40 p.

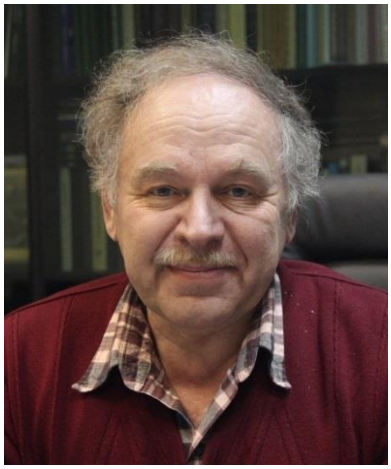
25. *Importing and Exporting Data with OJS.* URL: http://www.informatica.si/manual/ojs-import-export.pdf. URL: http://www.informatica.si/manual/ojs-import-export.pdf.

СВЕДЕНИЯ ОБ АВТОРАХ

ГЕРАСИМОВ Алексей Николаевич – аспирант Института математики и механики им. Н.И. Лобачевского Казанского (Приволжского) федерального университета.

Alex Nikolaevich GERASIMOV, received MS degree in Mathematics from Kazan Federal University, (2012). Currently is a graduate student at the N.I. Lobachevskii Institute of Mathematics and Mechanics of Kazan Federal University. Current scientific interests: data mining, knowledge extraction technologies, integration of scientific resources.

email: gerasimov.mailstore@gmail.com



ЕЛИЗАРОВ Александр Михайлович – доктор физико-математических наук, профессор, заслуженный деятель науки Республики Татарстан, зам. директора по научной деятельности Института математики и механики им. Н.И. Лобачевского Казанского (Приволжского) федерального университета.

Alexander Mikhailovich ELIZAROV – Doctor of Physics and Mathematics, Professor, Honoured Worker of Science of the Republic of Tatarstan, Deputy Director of the Lobachevskii Institute of Mathematics and Mechanics of Kazan Federal University. Current scientific interests: data mining, recommender systems, cloud computing, knowledge extraction technologies.

email: amelizarov@gmail.com

ЛИПАЧЁВ Евгений Константинович – кандидат физико-математических наук, доцент кафедры теории функций и приближений Института математики и механики им. Н.И. Лобачевского Казанского (Приволжского) федерального университета.

Evgeny Konstantinovich LIPACHEV – Candidate of Physics and Mathematics, Associate Professor, Lobachevskii Institute of Mathematics and Mechanics of Kazan Federal University. Current scientific interests: data mining, recommender systems, cloud computing, knowledge extraction technologies.

email: elipachev@gmail.com

Материал поступил в редакцию 15 января 2015 года