

# ИНФОРМАЦИОННЫЕ СИСТЕМЫ И СЕРВИСЫ КОМПЛЕКСНОЙ ПОДДЕРЖКИ ПЕРИОДИЧЕСКИХ НАУЧНЫХ ИЗДАНИЙ

Д. Ю. Ахметов<sup>1</sup>, А. М. Елизаров<sup>1</sup>, Е. К. Липачёв<sup>1</sup>

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

**Введение.** Сегодня общепризнано, что после получения нового научного знания и даже отдельного конкретного научного результата время, необходимое для их доведения до заинтересованного пользователя (ученого-теоретика, инженера-практика, менеджера, преподавателя, студента и других) должно быть максимально коротким. Именно по этой причине даже традиционные задачи, связанные с подготовкой научной работы и последующими процессами публикации статьи в научном журнале или материалах научной конференции и информирования заинтересованных читателей, решаются в настоящее время на новых организационном и информационном уровнях и базируются на широком использовании интернета и информационно-коммуникационных технологий (ИКТ). Кроме того, существенно изменилась инфраструктура современных научных изданий – речь уже идет не столько о формах и средствах использования ИКТ, сколько о создании программной платформы, реализующей развитую систему сервисов для работы с электронным контентом. При этом стираются различия между электронным и печатным изданиями, являющимися лишь различными формами представления соответствующей информации; без информационных сервисов и специализированных программ становится невозможным подготовить научную работу к публикации, а без сетевых коммуникаций – опубликовать ее. Учитывая, что в любом случае речь идет об использовании электронного контента, можно говорить об электронных научных изданиях, отдельно не выделяя печатные их формы, а электронные научные публикации можно рассматривать как разновидность электронных документов, особенности которых отражаются в их жизненном цикле.

Таким образом, современный электронный научный журнал должен базироваться на программной платформе, реализующей развитую систему сервисов для работы с электронным контентом. Отметим, что базовые понятия «электронный документ» и «жизненный цикл электронного документа», свойства электронных научных документов и их преимущества перед традиционными бумажными носителями обсуждены в [1].

В настоящей работе обсуждаются сервисы поддержки жизненного цикла научной публикации, определяющие функциональность современных программных платформ управления электронными научными журналами.

Особое внимание уделено сервису поддержки научного рецензирования – описан разработанный алгоритм автоматизации экспертной оценки, указаны изменения как в программном коде открытой системы Open Journal Systems, являющейся ядром программной платформы, так и в используемой базе данных.

**Жизненный цикл научной публикации и сервисы его поддержки.** Как известно, в сфере ИКТ для обозначения последовательности процессов от проектирования, разработки и сопровождения информационной системы до прекращения её эксплуатации используется термин «жизненный цикл». Каждый документ, в том числе электронный, за время своего существования также проходит определенные стадии, называемые «жизненным циклом документа». Обычно выделяют следующие его этапы: создание, описание, утверждение, хранение, поиск, извлечение и доставка. Жизненный цикл электронной научной публикации, в отличие от «бумажной», предполагает дальнейшее улучшение её качества и расширение свойств. Отметим в этой связи, например, новую технологию в области электронных научных публикаций, получившую название «живых документов» (см., например, [2, 3]). Как отмечено в [4], онлайн-представление становится центральным звеном публикационной деятельности ученого, и с него начинается жизненный цикл публикации.

Какими же должны быть сервисы поддержки жизненного цикла научной публикации? Анализ функциональности, реализуемой многочисленными современными программными платформами управления электронными научными журналами различных научных издательств, показал, что эта функциональность представлена рядом основных и дополнительных (опциональных) сервисов. В целом все их можно разделить на следующие группы: *сервисы подготовки научной публикации* – *сервисы для автора* (подготовки научной статьи в соответствии с редакционными требованиями, а также взаимодействия с редколлегией журнала), *сервисы поддержки научного рецензирования* (рецензента, редактора), *сервисы поддержки редакционных процессов* и *сервисы для читателя*. Все эти сервисы непосредственно связаны с процессами, реализуемыми с участием авторов и (или) редколлегии (редакции) журнала.

К основным процессам, связанным с деятельностью авторов и редколлегии (редакции) электронного научного журнала на начальном этапе, традиционно относят представление научной публикации, ее рецензирование и редактирование.

На этапе представления научной публикации важна автоматическая проверка статьи на соответствие стилевым правилам журнала (стилевая валидация), следовательно, нужна организация соответствующего автоматизированного сервиса подготовки научной публикации. В частности, в физико-математических статьях, набранных в T<sub>E</sub>X-нотации, должны использоваться стилевые файлы, специально подготовленные редакцией, и

соответствующие макрокоманды. Для физико-математических журналов необходима также программная поддержка процесса обработки электронных документов, содержащих формулы. Для работы с электронными документами, подготовленными в T<sub>E</sub>X-нотации, а это основной формат физико-математических журналов, разработан специализированный модуль стилевого оформления публикации по математике, размещенный на платформе science.tatarstan.ru (см. [5]). Соответствующий сервис выполнен в виде конструктора, предлагающего автору заполнить ряд полей, по которым автоматически формируется T<sub>E</sub>X-код статьи, выполненный по правилам журнала.

Другим сервисом для авторов является конструктор создания библиографических списков статей, учитывающий правила научного журнала и принятые сокращения названий журналов.

Еще одним сервисом для авторов, абсолютно необходимым при размещении информации о статье в цитатных базах данных (Scopus, Web of Science), является транслитерация библиографии публикации, т. е. процедура точной передачи в библиографическом списке статьи букв алфавита одного языка (и, соответственно, отдельных слов и текстов) средствами другой системы (алфавита другого языка), в частности, русских слов латиницей.

К сервисам поддержки научного рецензирования относится, прежде всего, автоматизация выбора рецензентов и самого процесса рецензирования. Отметим, что один из вариантов такой автоматизации реализован на платформе science.tatarstan.ru (см. [6]).

К редакционным процессам относят классификацию, аннотирование, выделение метаданных, публикацию, объединение в коллекцию, организацию поиска и навигации, долгосрочное хранение, конвертирование в различные форматы и распространение, контроль доступа, подписку, рассылку уведомлений, анализ статистики использования (включая формирование наукометрических данных). Сервисы, обеспечивающие автоматизацию перечисленных редакционных процессов, должны обеспечиваться любой информационной системой управления электронным научным журналом, их реализация на портале журнала, безусловно, необходима, однако набор таких сервисов должен постоянно расширяться в целях устойчивого развития издания. В частности, дополнительными сервисами поддержки редакционных процессов являются:

- проверка загружаемых текстов на плагиат;
- автоматизированная разметка статей, например, в математических электронных коллекциях (см. [7]);
- поиск и сбор OAI-метаданных (например, с помощью Open Harvester Systems);
- информетрический анализ (например, поддержка сервиса Article-Level Metrics (<http://article-level-metrics.plos.org>) в системе OJS);
- поддержка научных конференций (например, системы [www.easychair.org](http://www.easychair.org))

и Open Conference System);

- интеграция новостных лент научных журналов.

К набору сервисов для читателя можно отнести:

- оплату услуг (например, возможность работы с электронным кошельком PayPal ([www.paypal.com](http://www.paypal.com)));

- онлайн-общение (вебинары; видеоконференции);

- учет специфики обрабатываемых информационных ресурсов, например, расширенный поиск по фрагментам формул в математических коллекциях [7, 8], а также поиск на основе платформы семантической публикации математических документов [8, 9].

Итак, современные информационные системы управления электронными научными публикациями не должны ограничиваться сервисами удаленного представления статей в научный журнал и их дальнейшей обработки для окончательной публикации – в первую очередь необходимы доступ к сформированному контенту и расширенный поиск (по автору, названию статьи, ключевым словам и др.) в соответствующих электронных коллекциях. Тогда в полном объеме будут реализованы функциональные возможности, присущие электронным библиотекам.

**Программные платформы управления бизнес-процессами научного журнала.** Все названные выше функции и сервисы целесообразно обеспечить в рамках единой программной платформы. Внедрение такой платформы позволит, прежде всего, автоматизировать наиболее трудоемкие рабочие процессы, а порталное решение даст возможность интегрировать журнал в глобальное информационное научное пространство. Вместе с тем, для развития каждого электронного научного журнала необходимо постоянное расширение функциональности его базовой информационной системы.

В настоящее время успешное решение проблемы обеспечения оперативной доступности результатов научных исследований становится все теснее связанным с внедрением информационных издательских систем, разрабатываемых как в научных сообществах, так и в издательской отрасли в целом. Разработано и функционирует большое число информационных систем поддержки издательских процессов, которые делятся на системы с открытым и закрытым исходным кодом [10]. Последние принадлежат, как правило, ведущим мировым издательствам, обрабатывают по единым стандартам, с использованием единых технологий и единообразных требований тысячи журналов и не доступны небольшим издательствам (разве что на условиях аутсорсинга и, естественно, не бесплатно), выпускающим до нескольких десятков научных изданий. В такой ситуации небольшие издательства (в частности, издательства многих университетов и академических НИИ) при переходе в своей работе на современные ИКТ либо самостоятельно разрабатывают соответствующие информационные системы (которые, как правило, малоэффективны, не конкурентоспособны и, самое главное, не тиражируемы), либо используют системы с открытым кодом, дорабатывая их с

учетом собственных потребностей.

Среди систем управления электронными научными журналами с открытым кодом несомненным лидером стала программная платформа Open Journal Systems (OJS): согласно статистике OJS, в 2013 году система использовалась более чем в 6800 активных журналах, и это количество постоянно росло с момента начала ее внедрения (<https://pkp.sfu.ca/ojs/ojs-usage/ojs-map/>). Отметим, что преимущества использования систем с открытым кодом и выбора системы OJS подробно описаны в [10]. Там же предложена трехуровневая архитектура платформы управления электронными научными журналами, использующая в качестве ядра систему OJS.

**Технология расширения функционала OJS.** Система OJS имеет модульную архитектуру, что позволяет разрабатывать собственные классы и модули, расширяя ее функциональные возможности. OJS основана на паттерне Model-View-Controller (MVC), как следствие, хранилище данных, пользовательские интерфейсы и управляющие функции разделены на разные уровни взаимодействия. Система OJS платформонезависима и может быть установлена как под ОС Windows, так и на Unix-ориентированных операционных системах. В OJS используются свободно распространяемые язык программирования PHP, веб-сервер Apache, а также СУБД (MySQL, PostgreSQL).

Расширить функционал системы можно с помощью модулей, имеющих специальный формат (см., например, [11]). Большое количество инструментов, расширяющих возможности пользовательского интерфейса системы, а также обеспечивающих поддержку различных форматов данных, содержится в галерее модулей (плагинов) OJS (<http://pkp.sfu.ca/support/forum/viewforum.php?f=28>).

При самостоятельной разработке плагина для OJS необходимо вначале определить «тип» плагина, например, плагины, определяющие способы авторизации, относятся к типу «плагин авторизации», при разработке файлы плагина должны быть расположены в директории `\plugins\auth`. Плагины, не попадающие ни в один из типов, определяются как «общие» и должны быть расположены в директории `\plugins\generic`.

Как показал наш опыт, при разработке таких плагинов целесообразно следовать следующим правилам:

- все файлы создаваемого плагина должны быть размещены в отдельной папке;
- служебную информацию о плагине (дата создания, версия, тип, имя и др.) сохранять в xml-формате (например, в файле `version.xml`);
- код плагина подготовить на языке PHP и сохранить в файле `myPlugin.inc.php`, где `myPlugin` – название плагина;
- создать файл `index.php` для вызова плагина `myPlugin.inc.php`.

**Сервис поддержки научного рецензирования.** Важная отличительная особенность работы редколлегии научного журнала – наличие системы

независимого рецензирования и, как следствие, группы рецензентов-экспертов в предметных областях, определенных тематикой журнала. Процесс научного рецензирования – наиболее длительный этап в цепочке прохождения статьи в редакции научного журнала (от рукописи до публикации). В печатных научных журналах этот процесс мог затягиваться на годы. Критическими здесь, прежде всего, являются трудности в подборе рецензентов. Автоматизация этого процесса – выбор рецензента из базы экспертов с учетом не только компетенции эксперта, но и объема рецензирования и количества рецензируемых работ, – существенно сокращает сроки рецензирования. Отметим, что система автоматического назначения рецензента успешно использовалась в журнале *Lobachevskii Journal of Mathematics* [12].

Разработанный нами алгоритм автоматизации экспертной оценки (рецензирования) статьи, представленной в журнал, основан на использовании ключевых слов, которые в соответствии с установленными правилами должны быть указаны в статье, а также представлены автором при регистрации статьи в информационной системе научного журнала. С помощью ключевых слов, а также классификационных признаков (УДК, ББК, *Mathematics Subject Classification*) можно в автоматическом режиме достаточно точно определить область исследований, к которой относится представленная статья, и сопоставить эти данные с информацией об экспертах, содержащейся в базе данных информационной системы журнала [13, 14].

Решить задачу установления соответствия между множеством ключевых слов рассматриваемых научных работ и множеством, характеризующим специализацию рецензентов, можно различными способами. Наиболее простым (в плане технической реализации) является выделение одного из множеств в качестве базового – таковым целесообразно взять множество специализаций рецензентов и затем для каждой статьи определить несколько ключевых слов из этого множества. Форму отзыва рецензента целесообразно расширить за счет возможности добавления комментариев для выяснения мотивации проведенной им оценки.

Процесс экспертной оценки организован по стандартному принципу «принять в печать» – «вернуть с замечаниями на доработку» – «отклонить». В автоматическом режиме осуществляются контроль сроков рецензирования, а также рассылка уведомлений авторам и рецензентам. При выборе рецензентов система автоматически учитывает загруженность эксперта за счет ведения «шкалы загруженности» и истории рецензирования. Исключается возможность саморецензирования – эксперт, представивший статью в журнал в качестве одного из авторов, не сможет попасть в список возможных ее рецензентов.

Основные этапы алгоритма автоматического выбора экспертов научной статьи заключаются в следующем:

- формирование входных параметров – в качестве входного аргумента используется идентификатор научной работы (или первичный ключ сущности «научная работа» на языке баз данных);

- создание общего списка рецензентов – выборка осуществляется из базы данных рецензентов с учетом их специализации, загруженности, а также условия «рецензент не может быть автором»;
- отбор рецензентов по критерию наилучшего соответствия (большее число совпадений по ключевым словам, минимальные уровни загруженности, учет рейтинга рецензента и истории рецензирования);
- рассылка уведомлений рецензентам, авторам и редакторам.

Реализация описанного алгоритма потребовала изменений как в программном коде OJS, так и в используемой базе данных. Ограничения на возможность включения плагинов в систему OJS, связанные с использованием языка программирования PHP, потребовали изменения файла `peerReview.tpl`, а именно, к пункту меню «Выбрать рецензента» добавлен новый пункт меню «Automatic selection of the reviewer», перенаправляющий систему к плагину автоматизации. Параметры работы скрипта передаются методом GET, поэтому идентификатор статьи включается в адресную строку с указанием абсолютного пути к файлу скрипта:

```
<form>
  {assign var="idij" value=$submission->getId()}
  <a href='http://.../plugins/generic/autoReviewer/myPlugin.inc.php
    ?iu={$idij}'>Automatic selection of the reviewer</a>
</form>
```

После активации плагина считывается значение параметров из адресной строки и формируется список первичных ключей таблицы «roles», значения «user\_id» которых равно 4096 (т.к. данные обо всех пользователях OJS хранятся в одной таблице базы данных, каждая группа идентифицируется определенным натуральным числом, роль рецензента определяется числом 4096). Далее создаются массивы, состоящие из идентификаторов рецензентов и уровня их загруженности, производятся сортировка и отбор рецензентов. Формируются переменные «\$newdate» и «\$newdate2» для вычисления текущего времени и времени, к которому истекает срок подачи рецензии (по умолчанию OJS для составления рецензии выделяет 4 недели):

```
$newdate=date("Y-m-d H:i:s");
$newdate2=date ('Y-m-d 00:00:00', strtotime ('+4 weeks')).
```

В базе данных устанавливается связь между статьей и рецензентами, для этого в таблице «review\_assignments» создаются записи, количество которых отражает число отобранных рецензентов. В поля `submission_id` (идентификатор статьи), `reviewer_id` (идентификатор рецензента), `date_assigned` (текущее время) и `date_due` (конечная дата подачи рецензии) записываются значения '`$article`', '`$massiv_index[$index]`', `op'``$newdate`', '`$newdate2`'. После окончания всех операций с данными при помощи скрипта JavaScript организуется возврат в систему:

```
echo (" <script> location.href = 'http://.../editor/submissionReview/$article' </script>").
```

Вместе с тем, автоматический подбор рецензентов – только часть процесса автоматизации, необходимы также вспомогательные инструменты рецензирования, обеспечивающие загрузку наукометрических данных, поиск по статьям с близкой тематикой, сравнение полученных результатов с опубликованными ранее, в частности, на наличие заимствований и полноту ссылок.

Если научная статья прошла процесс рецензирования и получила положительные отзывы, то начинается редакционный процесс. На этом этапе происходят литературное редактирование статьи и ее перевод в форматы, поддерживаемые информационной системой журнала (pdf, html и др.). После этого рассматриваемая статья проходит техническое редактирование и при обнаружении ошибок вновь подвергается редакционному процессу. На данном этапе определяем также номер выпуска электронного издания, в котором статья будет опубликована. Процесс публикации непосредственно включает в себя компиляцию подготовленного электронного издания, устранение недочетов и технических неточностей.

**Особенности обработки математических текстов.** Большая часть электронных математических коллекций состоит из неструктурированных текстов в различных форматах. Ведущими мировыми издательствами физико-математической литературы, в том числе российскими, принята  $T_E X$ -нотация. Теговая система  $T_E X$ -документа позволяет придать ему структурную выразительность, недоступную другим системам нотации,  $T_E X$ -нотация обладает практически всеми необходимыми компонентами для создания семантической разметки, что особенно важно при создании процедур автоматической обработки математических документов электронной коллекции и использовании программ конвертации. Для работы с  $T_E X$ -документами функционал системы OJS необходимо расширить путем подключения специальных модулей, например, с помощью плагина, предложенного в [6].

Успешными примерами создания программной платформы управления математическими научными журналами служат широко известный портал Math-Net.Ru (<http://www.mathnet.ru/>) [15-17] и издательская платформа журнала Lobachevskii Journal of Mathematics [6, 12–14].

Работа выполнена при финансовой поддержке РФФИ (проекты 15-07-08522, 15-47-02472) и РГНФ (проект 14-03-12004).

## ЛИТЕРАТУРА

1. Елизаров А.М., Зуев Д.С., Липачёв Е.К. Управление жизненным циклом электронных публикаций в информационной системе научного журнала // Вестник Воронежского государственного университета. Сер. Системный анализ и информационные технологии. – 2014. – № 4. – С. 81–88.
2. Горбунов-Посадов М.М. Живая публикация // Открытые системы. – 2011. – № 4. – С. 48-49. Обновляемая («живая») публикация: URL: <http://keldysh.ru/gorbunov/live.htm>.



3. Паринов С.И., Когаловский М.Р. «Живые» документы в электронных библиотеках // Прикладная информатика. – 2009. – №6 (24). Авторская версия: <http://socionet.ru/publication.xml?h=repec:rus:isyigw:article-215>.
4. Горбунов-Посадов М.М. Интернет-активность как обязанность ученого // Информационные технологии и вычислительные системы. – 2007/ – № 3. – С. 88-93. – URL: <http://keldysh.ru/gorbunov/duty.htm>.
5. Ахметов Д.Ю., Елизаров А.М., Липачёв Е.К. Автоматизация процесса первичной обработки математической статьи в информационной системе электронного научного журнала// Тр. Математического центра имени Н.И. Лобачевского. Материалы Двенадцатой молодежной науч. шк.-конф. «Лобачевские чтения – 2013». – Казань: Изд-во Казан. матем. об-ва, 2013. – Т. 47. – С. 6-10.
6. Ахметов Д.Ю., Елизаров А.М., Липачёв Е.К. Система автоматизации редакционных процессов на платформе электронных научных журналов // Учёные записки Института социальных и гуманитарных знаний. – 2014. – №1 (12), Ч. 2. – С. 228-233.
7. Biryal'tsev E., Elizarov A., Zhil'tsov N., Lipachev E., Nevzorova O., Solov'ev V. Methods for analyzing semantic data of electronic collections in mathematics // Automatic Documentation and Mathematical Linguistics. – 2014. – V. 48, No 2. – P. 81-85.
8. Елизаров А.М., Липачев Е.К., Малахальцев М.А. Веб-технологии для математика. Основы MathML. – М.: Физматлит, 2010. – 192 с.
9. Елизаров А.М., Липачёв Е.К., Невзорова О.А., Соловьев В.Д. Методы и средства семантического структурирования электронных математических документов // Докл. РАН. – 2014. – Т. 457, № 6. – С. 642-645.
10. 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.
11. Willinsky J., Stranack K., Smecher A., MacGregor J. Open Journal Systems: a complete guide to online publishing. – Simon Fraser University Library, 2010. – 273 p.
12. 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.
13. Ахметов Д.Ю. Управление жизненным циклом электронной научной публикации // Тр. XV Всерос. науч. конф. «Электронные библиотеки: перспективные методы и технологии, электронные коллекции, RCDL-2013». Ярославль: Изд-во ЯрГУ им. П.И. Демидова, 2013. – С. 407-408.
14. Ахметов Д.Ю., Елизаров А.М., Липачёв Е.К. Модель сервисов электронного математического журнала и ее облачная реализация на платформе Open Journal Systems // Труды Российской школы «Математическое и

компьютерное моделирование фундаментальных объектов и явлений» и Международного научного семинара «Нелинейные поля в теории гравитации и космологии». – Казань: Отечество, 2013. – С. 86-92.

15. Жижченко А.Б., Изаак А.Д. Информационная система Math-Net.Ru. Применение современных технологий в научной работе математика // Успехи матем. наук. – 2007. – Т. 62, Вып. 5 (377). – С. 107-132.

16. Жижченко А.Б., Изаак А.Д. Информационная система Math-Net.Ru. Современное состояние и перспективы развития. Импакт-факторы российских математических журналов // Успехи матем. наук. – 2009. – Т. 64, Вып. 4 (388). – С. 195-204.

17. Chebukov D.E., Izaak A.D., Misyurina O.G., Pupyrev Y.A., Zhizhchenko A.B. Math-Net.Ru as a digital archive of the russian mathematical knowledge from the XIX century to today // Lecture Notes in Computer Science, V. 7961, ed. J. Carette et al., 2013. – P. 344-348, arXiv: 1305.5655