

С.Н.Верзунов

Институт автоматизации и информационных технологий НАН КР,

Бишкек, Кыргызстан

E-mail: verzunov@hotmail.com

ЭЛЕКТРОННЫЙ КАТАЛОГ ЖУРНАЛОВ НАУЧНОГО ИЗДАТЕЛЬСТВА НА БАЗЕ WORDPRESS

В настоящей работе предлагается решение задачи формирования информационного пространства научного издательства на основе системы управления веб-содержимым WordPress и программы разметки метаданных статей Artculus.

Ключевые слова: хранилище научной информации, обмен метаданными, плагин WordPress, Artculus, обработка XML-файлов.

Введение. Компьютеризация издательской деятельности в последние годы привела к существенному увеличению возможностей быстрой публикации результатов научных исследований с помощью различных издательских систем для публикации научных изданий в сети Интернет. Это либо коммерческие решения крупных издательств [1], либо различные самостоятельно разработанные системы, например, электронный научный журнал «Наука, новые технологии и инновации Кыргызстана» [2], или же решения на базе свободно распространяемого программного обеспечения [3].

Традиционно в Кыргызской Республике, в основном используются самостоятельные разработки или, реже, решения на основе свободно распространяемого программного обеспечения, так как большая масса журналов существует при поддержке небольших научных издательств и имеет ограниченные организационные и финансовые возможности. При этом компоновка цифровых научных коллекций происходит зачастую нерегулярно и зависит, в основном, от необходимости обеспечить информационную поддержку научной или образовательной деятельности. Это в частности относится и к сборникам научных статей, а также к материалам научных конференций. По большому счёту статьи из этих источников не имеют возможности попасть даже в такие хранилища научной информации, как Научная электронная библиотека (НЭБ) [4] или Киберленинка [5]. В качестве причин этого можно привести необходимость заключения лицензионного договора с авторами статей, а в основном, – нехватку человеческих и временных ресурсов. Кроме того, в большинстве из создаваемых электронных коллекций не только отсутствует возможность обмена метаданными научных статей, но и сами метаданные в каком-либо из общепринятых форматов их хранения. Это, в свою очередь, делает сомнительным возможность организации взаимодействия таких коллекций с основными хранилищами научной информации и создание условий для индексации ресурсов в ведущих системах поиска и хранения научных публикаций: НЭБ, Google Scholar, Web of Science и других [6]. Поэтому для подавляющего числа научных изданий в Кыргызской Республике актуальной целью является решение проблемы формирования информационного пространства, адекватного современным требованиям, предъявляемым к инновационной научной деятельности.

Одной из важных составляющих этого информационного пространства является интернет-сайт, обеспечивающий решение таких задач как: формирование веб-страниц статей внутри электронного каталога издательства, индексирующихся системами

поиска научных публикаций, создание инструментов экспорта статей в электронные коллекции. Особенно важной в настоящее время является проблема загрузки материалов из электронного каталога издательства в каталог НЭБ. В настоящей работе предлагается решение этой задачи путём разработки электронного каталога издательства на основе формата, предлагаемого самой НЭБ.

Формат хранения метаданных статей в Научной электронной библиотеке

НЭБ предлагает веб-программу разметки Articuluss [7] для подготовки выпусков журналов и неперiodических изданий в формате XML (eXtended Markup Language – расширяемый язык разметки). Структура XML-файла, генерируемого программой Articuluss, показана на рис. 1.

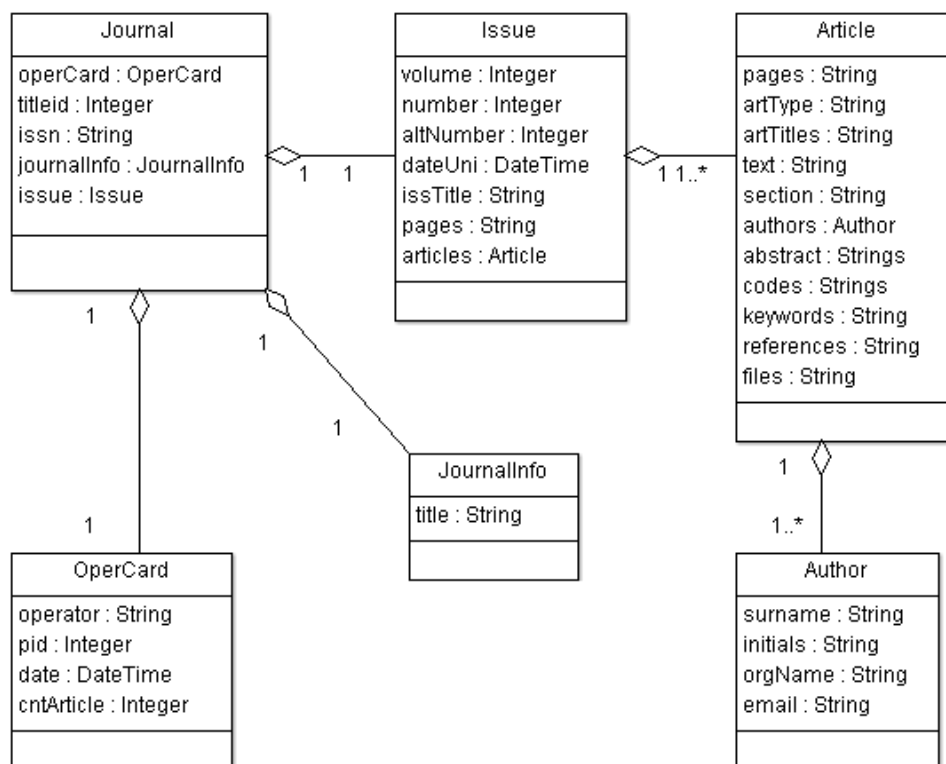


Рис. 1. Структура XML-файла, генерируемого программой Articuluss.

В основе этого формата лежит структура *Journal*, содержащая данные оператора (*operCard*), создавшего XML-файл, уникальный номер журнала в НЭБ (*titleid*), его ISSN, описание журнала в виде структуры *JournallInfo* и описание выпуска журнала в виде структуры *Issue*.

Структура *Journal* агрегирует несколько подструктур с разным уровнем вложения, в частности:

- *OperCard*, содержащую имя (*operator*), идентификатор (*pid*), дату регистрации (*date*), количество (*cntArticle*) созданных статей оператора, создавшего XML-файл;
- *JournallInfo*, содержащую название журнала (*title*);
- *Issue*, содержащую том (*volume*), порядковый номер (*number*), сквозной номер (*altNumber*), дату выпуска (*dateUni*), заголовок (*issTitle*), количество страниц (*pages*), и статьи выпуска (*articles*), заданные как набор структур *Article*.

Структура *Article*, в свою очередь, включает: диапазон страниц (*pages*), тип (*artType*), заголовки на русском и английском языках (*artTitles*), текст (*text*), раздел

(*section*), аннотацию (*abstract*), коды УДК, ГРТИ и др. (*codes*), ключевые слова (*keywords*), ссылки на цитируемую литературу (*references*), вложенные файлы (*files*) и одного или нескольких авторов (*authors*) статьи в виде набора структур *Author*. Структура *Author* в свою очередь содержит фамилию (*surname*), инициалы (*initials*), место работы (*orgName*) и электронную почту (*email*) каждого автора.

Хорошо продуманная сотрудниками НЭБ организация XML-файла, содержащего метаданные статей, позволяет оперативно организовать собственную систему обработки и отображения научных статей в электронном каталоге издательства. При этом полностью отпадает необходимость дважды выполнять разметку статей – сначала для электронного каталога издательства, а затем для НЭБ. И, что важнее, нет нужды создавать собственную программу для разметки статей, так как можно воспользоваться уже хорошо зарекомендовавшей себя программой Articulius. В качестве платформы для создания электронного каталога издательства предлагается использовать одну из наиболее широко распространённых [8] CMS (Content Management System – система управления веб-содержимым) – WordPress, а функции обработки и отображения статей организовать в виде подключаемого модуля, т.н. плагина.

Организация взаимодействия плагина с WordPress. Плагин WordPress – это набор функций, написанных на языке PHP и добавляющих определенные возможности для сайта, основанного на CMS WordPress, которые обычно объединяются с системой управления и функционалом WordPress при помощи Plugin Application Program Interface (PAPI) – программного интерфейса взаимодействия с плагинами. Разработанный плагин использует стандартный PAPI для взаимодействия с CMS WordPress, а именно: действия (*action*), фильтры (*filter*) и короткие коды (*shortcode*). В частности, короткий код *Metamaker*, необходимый для запуска плагина, должен быть размещен на главной странице каждого журнала в каталоге, фильтр *query_vars* – используется для управления плагином с помощью передачи параметров запроса пользователя, о чём будет сказано ниже. Кроме того применяется несколько служебных действий для загрузки каскадных таблиц стилей и java-скриптов (*admin_enqueue_scripts*, *wp_enqueue_scripts*), используемых для настройки визуального оформления плагина, и действие *admin_menu* для загрузки административного интерфейса управления плагином (рис. 2).

Таким образом PAPI, предоставляемый CMS WordPress позволяет легко расширить его функциональность без изменения исходного кода самого WordPress, за счёт достаточно просто подключаемых расширений, взаимодействующих с CMS посредством стандартного PAPI. Это предоставляет необходимые возможности для создания полнофункционального плагина, обеспечивающего загрузку, обработку и отображение статей из электронного каталога журналов научного издательства.

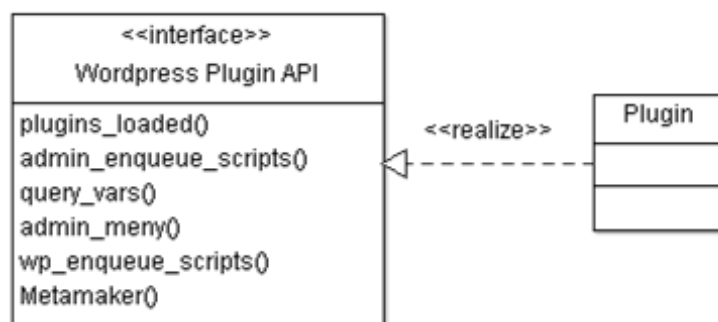


Рис. 2. Интерфейс взаимодействия плагина и CMS WordPress.

Основные функциональные возможности плагина. В целом все функции плагина (рис. 3) можно разделить на административные и пользовательские

возможности. Административные возможности плагина: выбор для редактирования необходимого журнала из каталога, просмотр имеющихся выпусков, добавление новых или удаление некоторых выпусков доступны только зарегистрированному администратору сайта. Пользовательские возможности: выбор для просмотра необходимого журнала, отображение списка выпусков, отображение содержания выпуска и сведений о выбранной статье доступны любому пользователю.

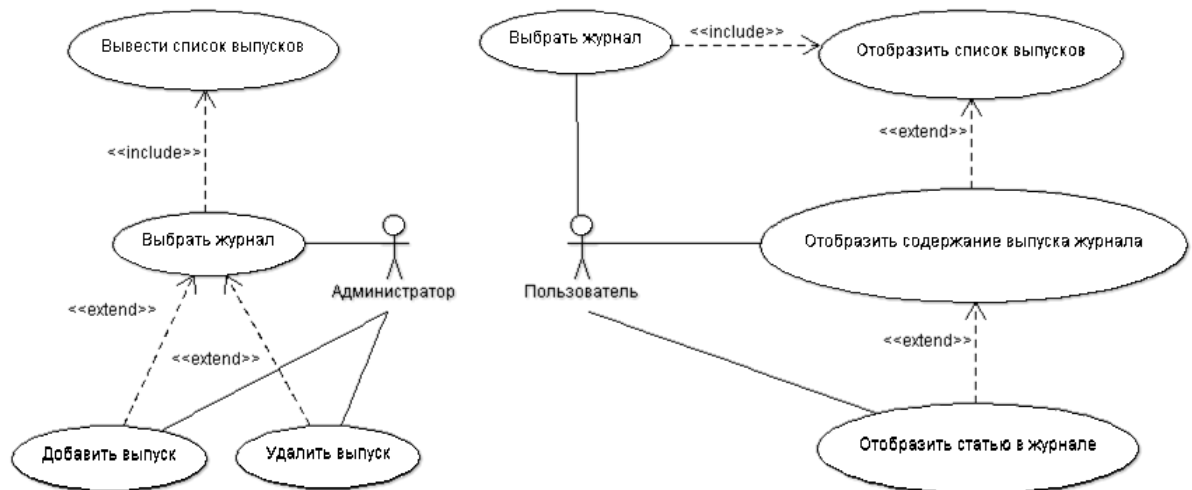


Рис. 3. Функциональные возможности плагина.

Эти возможности были реализованы в виде набора функций на языке PHP, взаимодействующих с одной стороны с CMS WordPress посредством описанного выше РАРІ, а с другой стороны – с пользователем, предоставляя ему простой и удобный web-интерфейс для просмотра электронного каталога журналов научного издательства.

Алгоритм и интерфейс взаимодействия плагина с пользователем. На основной странице каждого журнала (рис. 4) администратором сайта должен быть размещен короткий код *Metamaker*, обработав который, CMS WordPress запускает

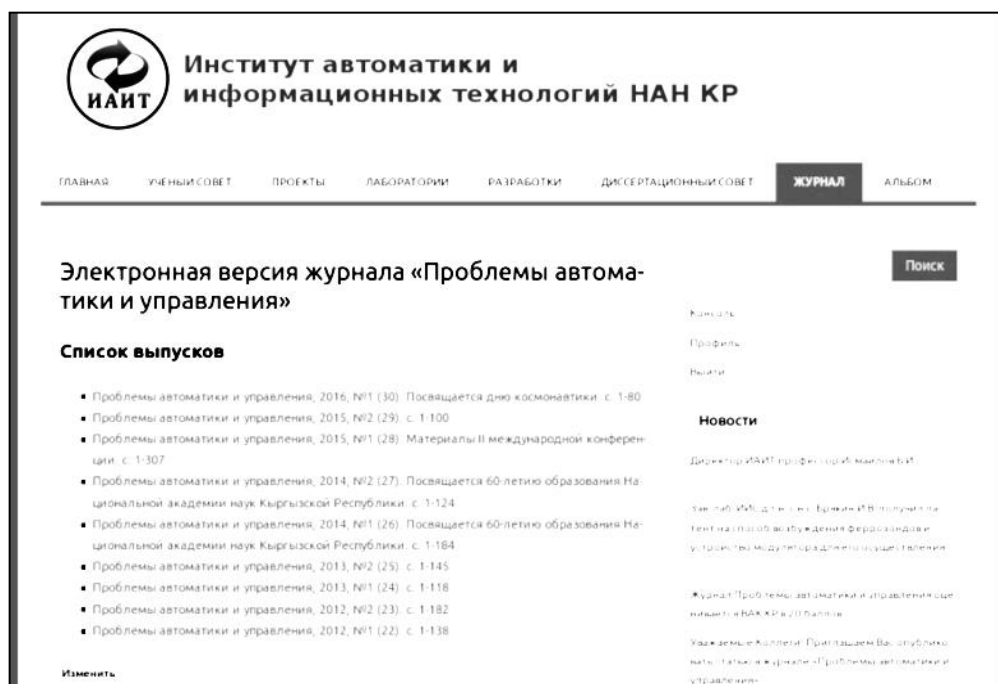


Рис. 4. Основная страница журнала.

алгоритм выполнения плагина (рис. 5). При этом каталог пакетов плагина проверяется на наличие XML- файлов. Затем по очереди плагином загружаются все найденные XML-файлы, содержащие выпуски журнала, на веб-страницу выводятся

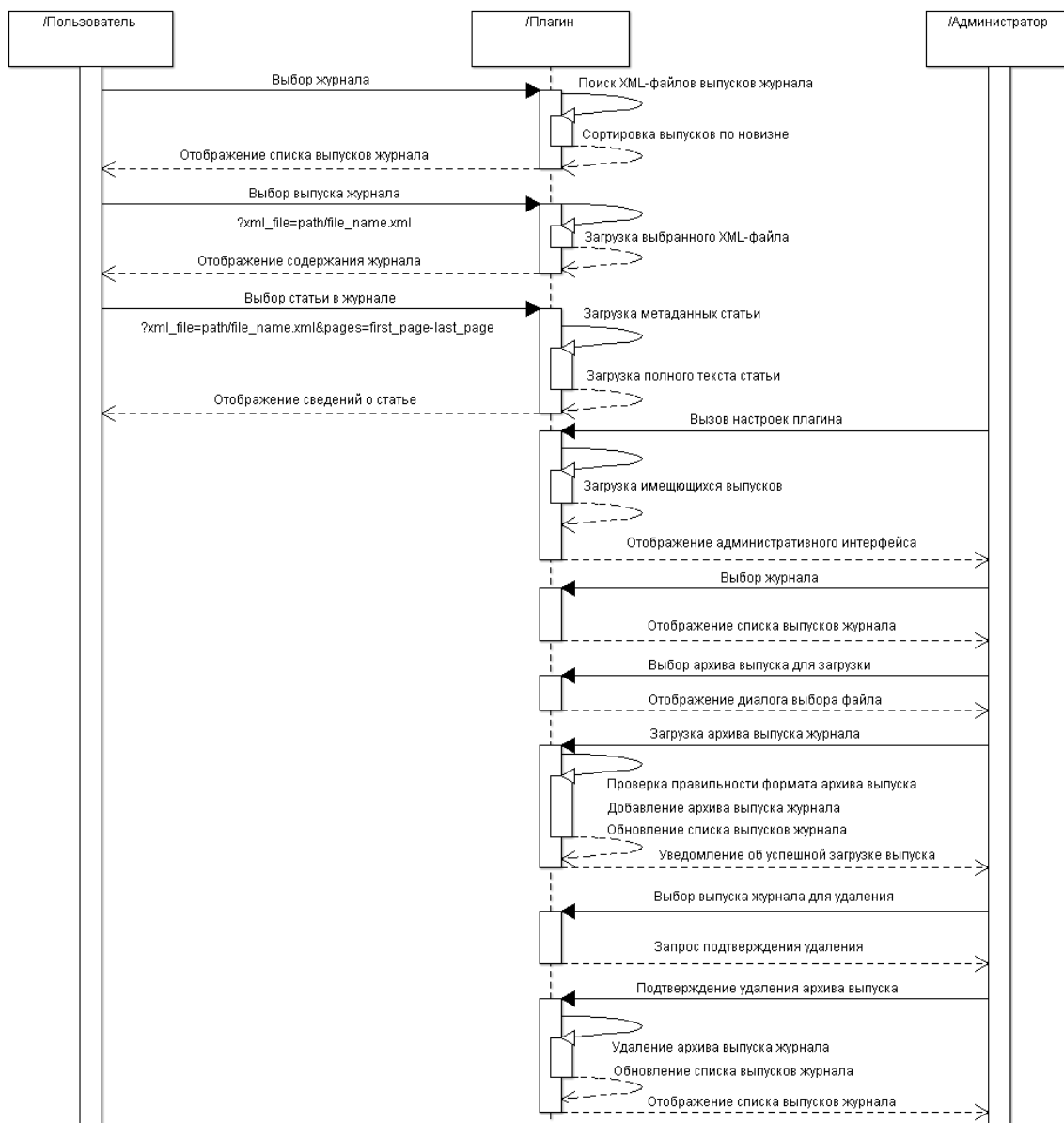


Рис. 5. Алгоритм работы плагина.

название журнала, год выпуска, номер и название выпуска, а также количество страниц в выпуске. При этом выпуски сортируются в порядке убывания сквозного номера журнала. Сведения о выпуске оформляются как гиперссылка, через которую пользователем осуществляется передача плагину имени XML-файла, содержащего выбранный пользователем выпуск, в формате:

?xml_file=path/file_name.xml,

где *path* – путь, и *file_name* – имя XML-файла.

После этого плагин загружает необходимый пользователю выпуск и отображает его содержание (рис. 6 а) с указанием названий статей, их авторов и номеров страниц статьи в выпуске журнала. Название статьи содержит гиперссылку следующего вида:

?xml_file=path/file_name.xml&pages=first_page-last_page,

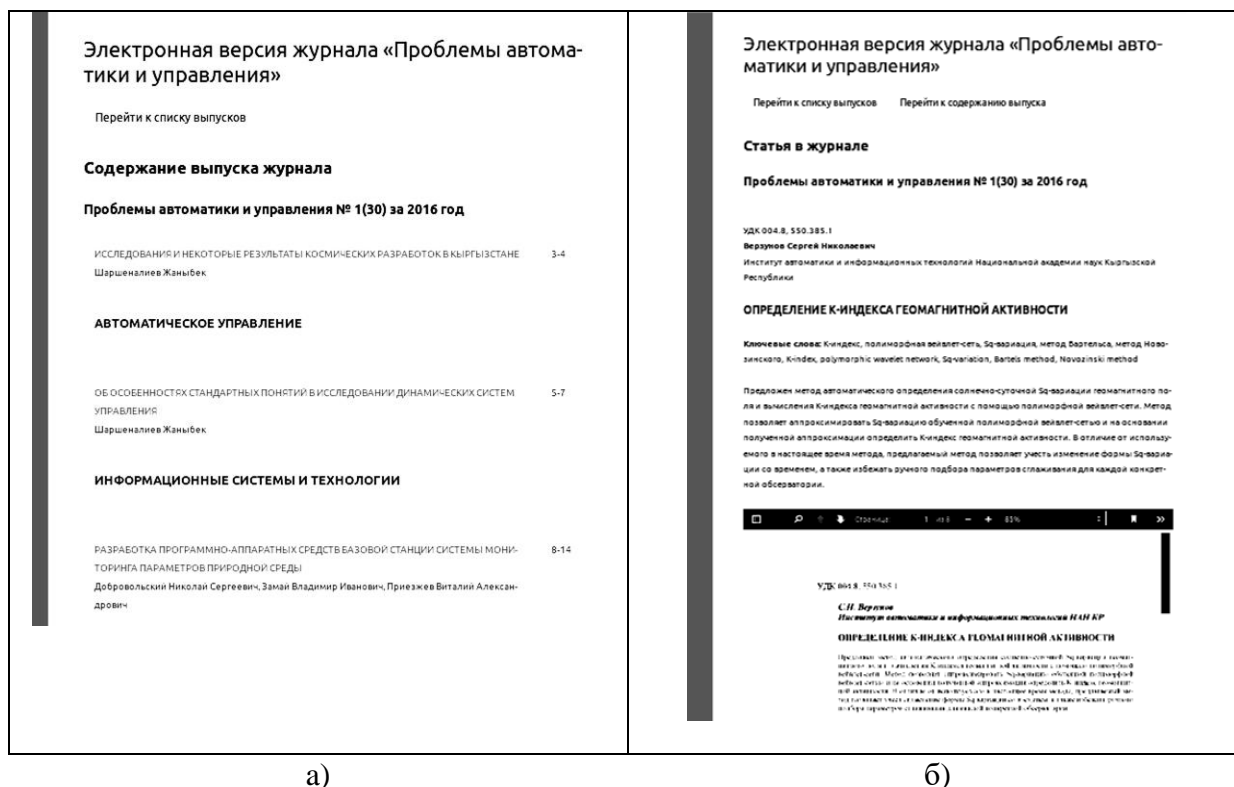


Рис. 6. Отображение плагином: а – содержания выпуска, б – статьи в журнале.

то есть, кроме имени необходимого XML-файла, она также содержит и номера страниц (*first_page-last_page*), соответствующие этой статье в выпуске журнала. Кроме того, на этой странице отображается кнопка «Перейти к списку выпусков», позволяющая вернуться к списку выпусков журнала.

При щелчке по интересующей пользователя ссылке на статью плагин загружает её содержание: УДК, фамилии, имена авторов, места их работы, название статьи, ключевые слова, аннотацию и полный текст статьи (рис. 6 б). Также на этой странице отображаются кнопки «Перейти к списку выпусков» и «Перейти к содержанию выпуска», позволяющие перейти к предыдущим страницам.

Административный интерфейс плагина (рис. 7) позволяет администратору сайта просмотреть список имеющихся выпусков, добавить новые или удалить некоторые.

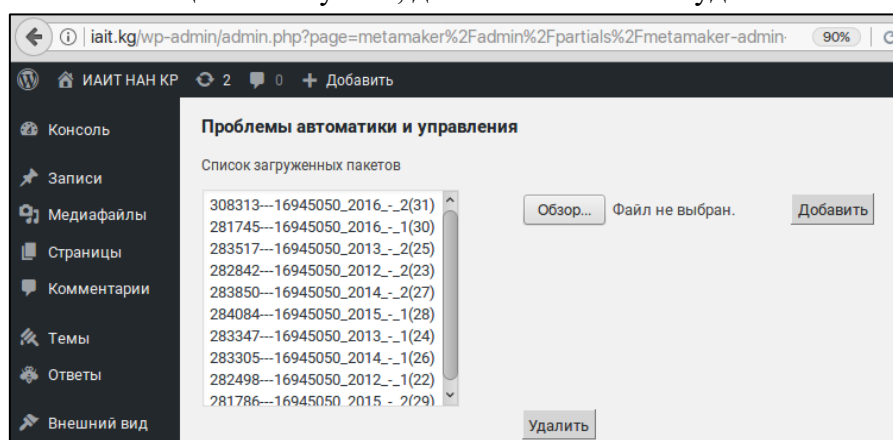


Рис. 7. Административный интерфейс плагина.

Выводы. Таким образом, разработан веб-сайт, содержащий электронный каталог журналов научного издательства, предоставляющий пользователям доступ к основной информации о каждом выпуске журнала, его содержанию, а также метаданным и полным текстам статей. Реализация этого функционала в виде подключаемого к широко распространенной в настоящее время системе управления веб-содержимым модуля предоставляет возможность достаточно оперативно организовать электронные каталоги журналов на базе уже существующих сайтов научных организаций. А использование в качестве основного формата хранения данных формата, принятого в Научной электронной библиотеке, позволяет значительно упростить разметку метаданных статей и их экспорт в различные репозитории научной информации. Разработанное программное обеспечение защищено свидетельством [9] Государственной службы интеллектуальной собственности и инноваций при правительстве Кыргызской Республики.

Литература

1. http://ranlib.ru/?page_id=14775 (дата обращения: 16.11.2016).
2. <http://195.38.182.113> (дата обращения: 16.11.2016).
3. http://www.isras.ru/index.php?page_id=1552 (дата обращения: 17.11.2016).
4. <http://elibrary.ru> (дата обращения: 17.11.2016).
5. <http://cyberleninka.ru> (дата обращения: 17.11.2016).
6. И. А. Мбого, Д. Е. Прокудин, А. В. Чугунов Комплексная интеграция цифровых коллекций в информационное пространство научных исследований // Технологии информационного общества в науке, образовании и культуре: Труды XVII Всероссийской объединенной конференции «Интернет и современное общество» (IMS-2014). – СПб: Университет ИТМО, 2014. – С. 48–53.
7. <http://intra.elibrary.ru/projects/articulus/help/articulus2.doc> (дата обращения: 18.11.2016).
8. <http://www.gs.seo-auditor.com.ru/cms/2016> (дата обращения: 19.11.2016).
9. Верзунов С.Н. Плагин WordPress для создания электронной библиотеки. (МЕТАМАКЕР) // Свидетельство Кыргызской Республики №443. 2016. 20160050.6 Бюл. № 1(213).