

---

---

ИНФОРМАЦИЯ

---

---

УДК 001.92:002.6:004.91

**ИНФОРМАЦИОННАЯ СИСТЕМА “АКУСТИКА”.  
РУССКОЯЗЫЧНЫЕ ИСТОЧНИКИ (HTTP://AKDATA.RU/)**

© 2015 г. В. Г. Шамаев\*, А. Б. Горшков\*\*, Н. В. Шамаев\*\*\*

\**Московский государственный университет им. М.В. Ломоносова, Физический факультет  
119992 ГСП-1, Москва, Ленинские горы  
E-mail: shamaev08@gmail.com*

\*\**Московский государственный университет им. М.В. Ломоносова,  
Государственный астрономический институт им. П.К. Штернберга*

\*\*\**Институт физики Земли им. О.Ю. Шмидта РАН*

Поступила в редакцию 22.12.2014 г.

Обсуждаются проблемы информационного поиска русскоязычных документов. Отсутствие этих работ в информационно-поисковых системах приводит к их потере для научного сообщества. Кратко описывается созданная “Информационная система “Акустика”. Русскоязычные источники”. На сегодня в нее загружено около 300 периодических изданий, книги и сборники трудов конференций за 2005–2014 гг. общим объемом около 30000 документов. В основном это все выходящие на русском языке печатные и электронные научные журналы РФ и стран СНГ, в которых встречаются материалы по акустике. Для основных журналов – тематических – глубина ретроспективы значительно больше. Обращается внимание на полнотекстовые сайты в Интернете журналов “Успехи физических наук” (<http://ufn.ru/>) и “Акустический журнал” (<http://www.akzh.ru/>), а также на еще один сайт – “Сигнальная информация. “Акустика”” (<http://akinfo.ru/>), на котором публикуется текущая информация по акустике.

*Ключевые слова:* интернет-ресурс, информационно-поисковые системы, сигнальная информация по акустике, русскоязычные источники, электронные документы, базы данных.

DOI: 10.7868/S0320791915030156

Уже неоднократно отмечалось [1–3, 10], что “поисковики” в Интернете по ключевым словам: Google, Yahoo!, DuckDuckGo и др., оригинальные российские поисковые машины Яндекс, Рамблер, Mail.Ru предоставляют огромное количество релевантной информации, но ее пертинентность, т.е. соответствие цели пользователя, заключающейся в поиске нужной информации, весьма низка. Можно сказать, что такое случается из-за неудачно составленного запроса, что, в общем-то, верно, и путем многочисленных приближений можно в конце концов достичь приемлемого результата. Но при этом в выдаче получаем, как правило, уже то, что и так нам известно. Теряем информацию, о которой не знаем. В качестве примера предлагаем поискать в Google по ключевому слову “акустика”, и вас завалят с первой же страницы рекламными текстами об акустических концертах, акустических системах, сабвуферах, наушниках Sony, Toshiba, JBL и т.п. Ранее функцию по выборке и фильтрации информации довольно успешно выполняли печатные реферативные журналы (РЖ), которые по разным причи-

нам прекратили свое существование [4–6]. В первую очередь – из-за невозможности листать тысячи страниц в поисках нужной информации, так как ее объемы на порядки увеличились по сравнению с объемами, что существовали при возникновении РЖ. Не последнюю роль в закрытии для некоторых из них сыграла и низкая оперативность отражения материала, что и раньше-то было неудобно, а сейчас просто неприемлемо; также почти решающее влияние собственно на прекращение их выпуска внесли огромные полиграфические затраты. Чтобы избежать информационного коллапса, стали создаваться электронные тематические базы данных научной информации, в которых присутствует оперативность, нет расходов на печать и в то же время отсутствует паразитная информация и фильтруется информационный шум, присущие вышеназванным поисковикам.

Можно ностальгировать о старом добром РЖ ВИНТИ СССР [1], наполненном добротной информацией, который поступал к зарубежному читателю по подписке, многотомном Chemical

Abstracts Services [6], отражавшем и русскоязычные источники, или томах в твердой белой обложке Astronomy and Astrophysical Abstracts, но таковы реалии нашего времени — времени цифровой эпохи и отсутствия сколько-нибудь существенной государственной поддержки научных исследований. Вместе с закончившейся эрой печатных РЖ пришла в упадок технология централизованного сбора, переработки и хранения научно-технической информации, ее анализ, которыми, в частности, славилась реферативная служба ВИНТИ СССР.

С развитием электронных средств предоставления информационных услуг большое значение приобретает такой параметр как скорость обработки научной информации в информационных центрах и время доставки ее потребителю. Так, например, в реферативной службе ВИНТИ РАН временной промежуток от поступления единицы информации (книга, журнал, депонированная рукопись, патент и т.д.) до отражения ее в РЖ ВИНТИ в 1990-е годы увеличился до 6–10 мес., а зачастую и более [1, 3–5, 11–12]. Это связано с долгим процессом обработки документов на входе, переработки их в отделах научной информации при реферировании и редактировании, неторопливой загрузки в БД. Как ни странно, это относится и к русскоязычным источникам, которые, как нам казалось, должны обрабатываться быстрее, полнее и качественнее, так как при них имеется русскоязычный “реферат” — резюме [7–9]. К тому же русскоязычные источники весьма слабо представлены, что ниже демонстрируется на примере выпуска “Акустика” РЖ “Физика” ВИНТИ. В “западных” информационно-поисковых системах (ИПС), хотя они и оперативны, русскоязычные источники теперь совсем не представлены и, следовательно, недоступны зарубежному читателю за исключением журналов, которые переводятся и издаются зарубежными издательствами [2].

Информационное обеспечение стало личным делом каждого научного работника или технического специалиста. К счастью, появление персональных компьютеров и Интернета несколько сгладило возникшую проблему, но вовсе не решило. Бурный рост научно-технической информации давно уже не позволяет отдельному человеку успешно справляться с ее интеллектуальной переработкой. Во всем мире давно начали появляться базы данных и основанные на них информационно-поисковые системы, создаются тематические базы данных научной информации. Но это в мировом масштабе и относится в основном к англоязычной информации.

Что же касается русскоязычной информации, то периодические издания хорошо представлены в “Научной электронной библиотеке eLIBRARY.RU” и по соответствующей тематике на Общероссийском математическом портале Math-Net.Ru. Мы даже не будем говорить о пропусках в их наполнении, каком-

то количестве опечаток, недостатках, на наш взгляд, поискового аппарата — без этого не обойтись, но это все мелочи по сравнению с глобальностью этих проектов и их удачным воплощением.

Целью же нашей работы было предложить информационный продукт по русскоязычным работам, который максимально полно может отобразить какую-либо тематическую область, и более “повернут” к требованиям специалистов, а именно, кроме поиска по источникам, периодам, авторам, ключевым словам, имеет рубрикатор — все документы прорубрицированы, и поиск по рубрикам входит в общую систему поиска. Отсутствие рубрикатора и, следовательно, рубрицирования документов на самом деле превращает БД даже с хорошим наполнением просто в хранилище документов, и никакой поиск даже по ключевым словам не поможет нахождению необходимой новой информации. В связи с этим ключевое значение приобретает качество проведенного рубрицирования. Поставленная не на ту полку книга в библиотеке все равно, что потеряна.

Важным качеством подобных тематических ИПС должно быть их постоянное пополнение как текущими, так и архивными материалами и возможность отслеживания интересов пользователей. Кроме простого предоставления информации, на их основе могут проводиться работы по изучению состояния науки на текущий момент времени, сниматься данные для последующего анализа тенденций в развитии той или иной области [1–2, 8, 10].

Создание ИПС — не только огромный, наполненный информацией объем, это видимый результат, в создании которого присутствует и научная работа. Результаты ее в виде аналитико-синтетической переработки информации дают возможность многочисленным исследователям проводить анализ, сопоставление и проверку возникающих гипотез, синтезировать новое знание.

На основе “западных” БД были созданы, стали доступны и пользуются популярностью у нас в стране такие ИПС как SciFinder, Web Of Science, База рефератов ADS (NASA), Scopus, Google Scholar и др. В то же время, полноценной ИПС по русскоязычным источникам до сих пор нет. Такой, уникальной по русскоязычному наполнению системой могла бы стать БД ВИНТИ РАН с ее ИПС “Сокол”, но ее исполнение а, главное, наполнение оставляют желать лучшего [1, 3–5, 7]. Попробуйте обратиться к ней на сайте ВИНТИ РАН — <http://www2.viniti.ru/>.

Вот на таком фоне и начала создаваться ИПС “Акустика” по русскоязычным источникам (<http://akdata.ru/>) — такая специализация вызвана, как уже отмечалось, почти полным отсутствием источников нашей страны и других стран СНГ в западных ИПС.

Рис. 1. Информационная система “Акустика”. Русскоязычные источники. Интерфейс пользователя.

Портал “Информационная система “Акустика”. Русскоязычные источники” (<http://akdata.ru/>) является первым этапом создания ИПС по физико-математическим наукам. В настоящее время решены технологические вопросы и функционирует сама ИПС “Акустика”. Наполнение БД системы реализовано со стадии сбора, рубрицирования и редактирования материалов, экспорта документов в БД и до представления ее в Сети с соответствующим интерфейсом пользователя. На портале имеются также точки входа на сайты “Акустического журнала” [8] и “Сигнальной информации” [9]. Авторы имеют длительную историю создания информационно-поисковых систем в ВИНТИ РАН [4–5, 11–12], но получили возможность разработки такого продукта только на кафедре акустики физического факультета МГУ им. М.В. Ломоносова. Именно поэтому ИПС первоначально ориентирована на акустику.

В концепцию системы заложены следующие принципы:

- привлечение максимально возможного количества источников информации (периодические издания, книги, сборники трудов конференций и т.д.);
- разработка специального рубрикатора и рубрицирование каждого документа;
- создание структур основной БД и Интернет-БД;

- разработка интерфейса пользователя;
- привлечение максимально возможного количества пользователей к подготовке материала по своим работам. Это помогает более тщательному рубрицированию материала.

Работа велась в трех направлениях:

- получение информационного ресурса в результате поиска, структурирования и аналитико-синтетической обработки информации;
- разработка новых и развитие существующих программных средств, использующихся для обработки, хранения, получения и распространения информационных продуктов;
- разработка средств веб-доступа к полученным информационным продуктам.

ИПС “Акустика” (рис. 1, 2) на сегодня отражает около 300 периодических изданий, книги и сборники трудов конференций за 2005–2014 гг. и в основном покрывает все выходящие на русском языке как печатные, так и электронные научные журналы РФ и стран СНГ. По “полноразметочным” журналам наполнение составляет весь доступный нам период их выхода. Поиск возможен по источникам (названиям, томам, номерам), году (годам) выхода, авторам, ключевым словам, рубрикам.

Проведенное сравнение с поиском в Google Scholar (<http://scholar.google.ru/>) показывает, что по взятым произвольно источникам, авторам, ключевым словам в начальных страницах его вы-

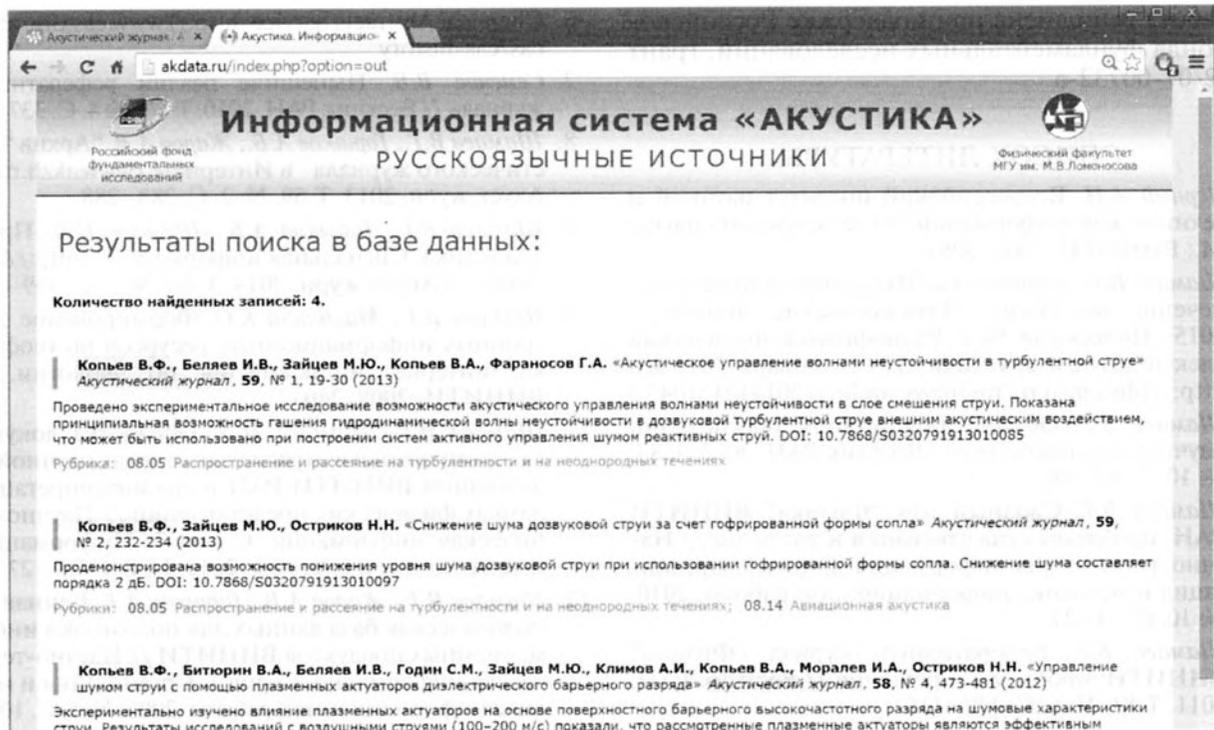


Рис. 2. Информационная система «Акустика». Русскоязычные источники. Вид окна «Результаты поиска».

дачи присутствует только малая часть документов, выданных ИПС «Акустика». Можно было бы провести сравнение с ИПС ВИНТИ РАН, если бы она была доступна. Но это не столь важно. Эта ИПС отражает содержание выпуска «Акустика» РЖ «Физика» ВИНТИ, но в этом выпуске из года в год, начиная с 1990-х гг. отражается лишь малая часть русскоязычных документов [8, 9], мы не говорим уже об оперативности: в 2010 г. – 187 документов, в 2012 – 254 и, после критики [3–5, 8], в 2013 – 581; 2014 год еще не закончен, но он на этом же уровне (рис. 3). На рисунке пропущен 2011 год как нетипичный: в течение этого года нами проводился эксперимент по наполнению выпуска «Акустика» русскоязычными документами, и их количество превысило 3500 за год, что подтвердило необходимость разработки ИПС.

В заключение отметим, что популярностью пользуются все хорошо организованные ИПС с удобным интерфейсом. Мы уже говорили о портале «Научная электронная библиотека eLIBRARY.RU» и Общероссийском математическом портале Math-Net.Ru. К популярным ИПС, несомненно, относится сайт журнала «Успехи физических наук» (<http://ufn.ru/>): хороший интерфейс, удобный поиск, возможность прочитать каждую статью, оперативность отражения текущего выпуска и, конечно, свободный доступ. К ним же можно отнести сайт «Акустического журнала» (<http://www.akzh.ru/>), который создан в 2012 г. по

инициативе академика О.В. Руденко, а также, надемся, нашу «Сигнальную информацию по акустике» (<http://akinfo.ru/>) [9]. Образно можно сказать, что последняя следует технологии «вики» (от гавайского wiki-wiki – «как можно быстрее»), непрерывно отображая текущую информацию из более чем 300 периодических русскоязычных научных журналов. Надеемся, что и ИПС «Акустика» будет развиваться и пользоваться популярностью, ведь количество научных сотрудников и специалистов в области акустики не меньше, чем все количество физиков в нашей стране. Акустикой занимаются биологи, геологи, гидрологи, музыканты, океанологи, строители, экологи и т.д.

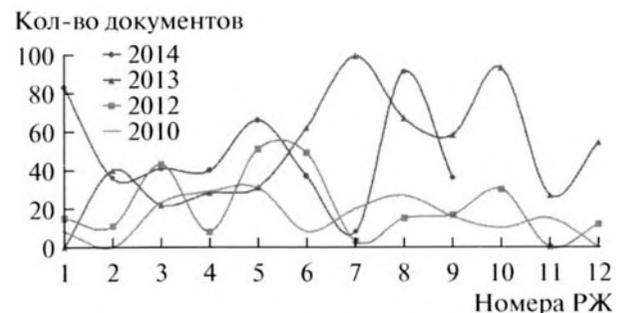


Рис. 3. Количество русскоязычных документов в выпуске «Акустика» РЖ «Физика» ВИНТИ за 2010, 2012–2014 гг.

Работа выполнена при поддержке Российского фонда фундаментальных исследований, грант № 12-07-00732-а.

### СПИСОК ЛИТЕРАТУРЫ

1. *Черный А.И.* Всероссийский институт научной и технической информации: 50 лет служения науке. М.: ВИНТИ, 2005. 306 с.
2. *Шамаев В.Г., Горшков А.Б.* Информационное обеспечение акустики. “Ломоносовские чтения — 2015” Подсекция № 2. Радиоп физика, физическая электроника и акустика (13–16 января 2015 г.). <http://phys.msu.ru/gus/news/archive/201412121045/>
3. *Шамаев В.Г.* Об информационном обеспечении научных исследований // Вестник РАН. 2013. Т. 83. № 10. С. 43–48.
4. *Шамаев В.Г.* Сводный том “Физика” ВИНТИ РАН: проблемы существования и развития // Научно-техническая информация. Серия 1: Организация и методика информационной работы. 2010. № 10. С. 21–27.
5. *Шамаев В.Г.* Реферативный журнал “Физика” ВИНТИ: проблемы и решения // Вестник РАН. 2011. Т. 81. № 5. С. 430–435.
6. Chemical Abstracts Service. <http://www.cas.org/about-cas/cas-history>.
7. *Семенов В.В.* Нынешние реалии реферативного журнала // Вестник РАН. 2010. Т. 80. № 4. С. 337–341.
8. *Шамаев В.Г., Горшков А.Б., Жаров А.В.* “Архив “Акустического журнала” в Интернете ([www.akzh.ru](http://www.akzh.ru))” // Акуст. журн. 2013. Т. 59. № 2. С. 283–288.
9. *Шамаев В.Г., Горшков А.Б., Шамаев Н.В.* Проект “Акустика. Сигнальная информация” (<http://akinfo.ru/>) // Акуст. журн. 2014. Т. 60. № 1. С. 109–114.
10. *Шамаев В.Г., Малинина К.О.* Формирование электронных информационных ресурсов по геофизике: интернет-энциклопедия по экологии. М.: ВИНТИ, 2008. 240 с.
11. *Шамаев В.Г., Шамаев Н.В.* Обработка документального потока в отраслевых отделах научной информации ВИНТИ РАН и его интерпретация в рамках физических представлений // Научно-техническая информация. Серия 2: Информационные процессы и системы. 2010. № 8. С. 19–27.
12. *Шамаев В.Г., Жаров А.В., Горшков А.Б.* Единая технологическая база данных для подготовки информационных продуктов ВИНТИ // Научно-техническая информация. Серия 1: Организация и методика информационной работы. 2006. № 5. С. 10–15.