

ИПС в Сети и информационное обеспечение акустикиВ. Г. Шамаев^{1,*}, А. Б. Горшков², Л. Г. Гущина³, В. И. Якименко⁴¹Московский государственный университет имени М. В. Ломоносова, физический факультет, кафедра акустики, Россия, 119991, Москва, Ленинские горы, д. 1, стр. 2²Московский государственный университет имени М. В. Ломоносова, Государственный астрономический институт имени П. К. Штернберга, Россия, 119991, Москва, Университетский пр-т, д. 13³Калужский филиал Московского государственного технического университета имени Н. Э. Баумана, Россия, 248000, Калуга, ул. Баженова, д. 2⁴Тюменский государственный университет, Россия, 625003, Тюмень, ул. Семакова, д. 10
(Статья поступила 09.02.2015; Подписана в печать 18.02.2015)

Обсуждаются проблемы поиска русскоязычных работ информационно-поисковыми системами (ИПС). Отсутствие этих работ в зарубежных базах данных приводит к их потере для научного сообщества. В связи с этим обращается внимание на русскоязычные ИПС — Научную электронную библиотеку (<http://elibrary.ru/>), портал MathNet (<http://www.mathnet.ru/>) и критикуется русскоязычная часть по акустике АБнД ВИНТИ РАН (<http://www2.viniti.ru/>). Кратко описывается созданная «Информационная система «Акустика». Русскоязычные источники» (<http://akdata.ru/>). В наполнении системы используются около 300 периодических изданий России и стран СНГ, а также книги и труды конференций. Особое внимание уделено полнотекстовым архивам «Успехов физических наук» (<http://ufn.ru/>) и «Акустического журнала» (<http://www.akzh.ru/>), а также сайту «Сигнальной информации «Акустика»» (<http://akinfo.ru/>). Проекты, которые имеют узкотематическую направленность, не смогут заменить зарубежные информационно-поисковые системы, отражающие десятки тысяч научных журналов (Web of Science, SCOPUS, SciFinder, INSPEC и др.). Но, созданные специалистами и для специалистов, даже небольшие по временному промежутку, но полные по охвату источников ресурсы представляют огромную ценность и востребованы научным сообществом. В дальнейшем они могут быть объединены в большие тематические БД.

PACS: 43.10.Pr

УДК: 001.92; 002.6; 004.91

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

ВВЕДЕНИЕ

С начала 1990-х гг. в России, да и в странах бывшего СССР задачи информационного обеспечения науки и техники перестали быть в числе приоритетных и хорошо финансируемых, отошли на второй план. На порядок и более уменьшились тиражи научных журналов, резко сократились расходы как на глобальную информационную службу — ВИНТИ РАН с её Реферативным журналом (РЖ), так и на отраслевые — центры научно-технической информации и научные библиотеки. Глубину кризиса хорошо демонстрирует упадок РЖ как по наполнению и оперативности, так и по тиражам. Приведем данные по РЖ ВИНТИ [1, 2] (тиражи — рис. 1, наполнение — рис. 2).

Таким образом, как мы уже неоднократно отмечали, закончилась эра печатного РЖ ВИНТИ, имевшего на протяжении 1950–1980-х гг. многотысячную подписку по большинству его тематических фрагментов, а вместе с ней пришла в упадок технология централизованного сбора, обработки и хранения научно-технической информации, её анализ, которыми славился ВИНТИ СССР. И закончилась не сегодня, а в начале 1990-х гг.

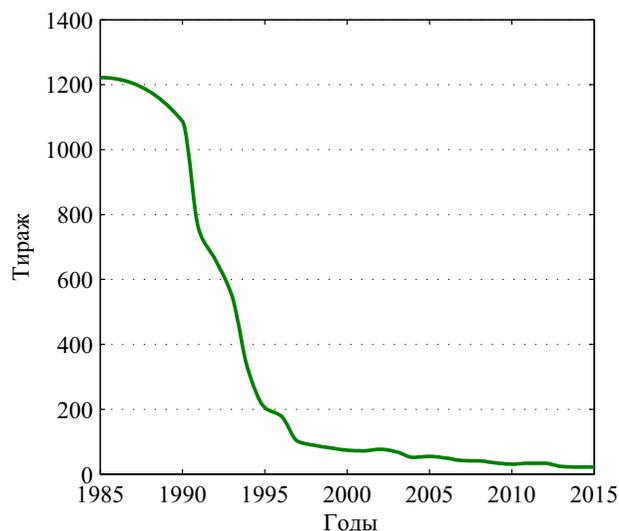


Рис. 1: Тираж РЖ «Физика» за 1985–2015 гг.

Мы этого как-то не заметили, но Реферативный журнал обвалился сразу (рис. 2), потеряв и наполнение, и оперативность, и охват источников, и, главное, хороших референтов и редакторов [2–4, 12]. Информационное обеспечение стало личным делом каждого науч-

*E-mail: shamaev08@gmail.com

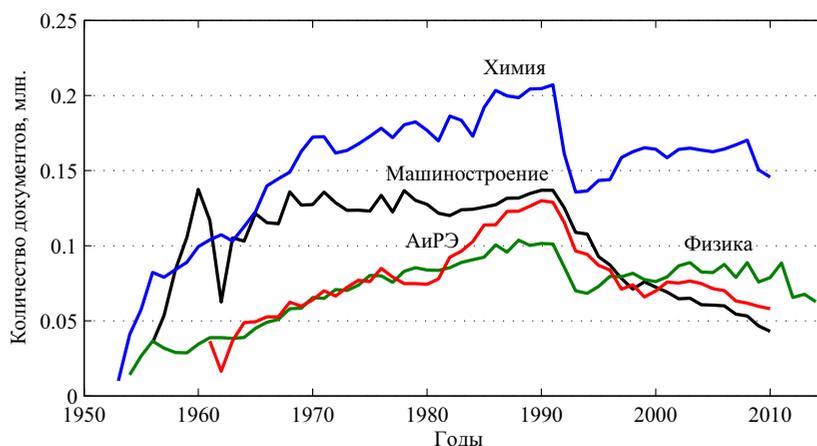


Рис. 2: Наполнение выпусков «Машиностроение», «Химия», «Физика», «Автоматика и радиоэлектроника» РЖ ВИНТИ (1953–2014 гг.)

ного работника или технического специалиста. Можно сожалеть об этом, но это судьба всех печатных РЖ в мире. Она не обошла ни Chemical Abstracts, ни Physics Abstracts, ни близкий нам Astronomy and Astrophysics Abstracts, ни другие, подобные им печатные издания. Многие из них заблаговременно были переведены в электронный вид.

1. ИНФОРМАЦИОННО-ПОИСКОВЫЕ СИСТЕМЫ В ИНТЕРНЕТЕ

Для некоторого количества научных работников проблема информационного обеспечения своих исследований не стоит. Есть Интернет, есть несколько научных журналов по тематике, и этого достаточно. Но что хотел бы продвинутый пользователь:

1. возможность постоянно иметь доступ к свежей информации;
2. возможность получать полную информацию, т. е. иметь доступ и к архивной информации;
3. возможность легко находить информацию;
4. возможность бесплатного доступа и т. д.

В докомпьютерную эру это всё худо-бедно обеспечивали научные библиотеки. Они централизованным образом хорошо комплектовались источниками, был достаточно большой штат сотрудников.

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 справиться с её интеллектуальной переработкой. В первую очередь, из-за невозможности листать тысячи страниц многочисленных печатных изданий в поисках нужной информации, т.к. их объемы на порядки

увеличились по сравнению с объемами, что существовали при их возникновении.

Докомпьютерная эра ушла. Сейчас другое время. Есть Сеть. Однако уже неоднократно отмечалось, 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. Также и поиск в них связан с рядом неудобств:

1. огромным информационным «шумом», засильем паразитной информации;
2. слабой структурированностью информации;
3. многократным дублированием информации;
4. отсутствием редактирования информации и, в силу этого, наличием большого количества ложной информации и т. п.

Притом, только Google Scholar имеет доступ к статьям из научных журналов практически всех наиболее престижных издательств.

Кроме того, как многие неожиданно для себя выяснили, большое значение имеет проблема правильного составления поискового запроса. Ранее функцию по выборке и фильтрации информации довольно успешно выполняли печатные реферативные журналы, но, как мы уже отмечали, они по разным причинам прекратили свое существование. Не последнюю роль в закрытии для некоторых из них сыграла плохая наполняемость и низкая оперативность отражения материала, что и раньше-то было неудобно, а сейчас просто неприемлемо. Почти решающее влияние собственно на прекращение их выпуска внесли огромные полиграфические затраты. Чтобы избежать информацион-

ного коллапса, стали создаваться электронные политематические и тематические базы данных (БД) научной информации, в лучших из которых присутствует оперативность, нет расходов на печать (документооборот происходит в электронном виде), проводится квалифицированное редактирование и, поэтому, отсутствует паразитная информация и фильтруется информационный шум, присущие вышеназванным поисковикам.

Это основные причины появления специализированных информационных систем — как глобальных, охватывающих множество тематик, так и локальных, посвященных узким вопросам. Их популярность связана с хорошим наполнением и удобным интерфейсом для пользователя и, как правило, оперативностью. Так, Chemical Abstracts Service (CAS) тратит на обработку источников не более одного месяца с момента их поступления.

Благодаря грантам РФФИ на информационные ресурсы, стали доступны в России и пользуются популярностью у продвинутой части пользователей такие информационно-поисковые системы как Web of Science (Thomson Reuters), SciFinder (Chemical Abstracts Service). Министерство образования и науки профинансировало доступ к INSPEC и SCOPUS.

В этих базах данных мы почти всегда найдем то, что ищем (книгу, журнал, статью), но практически никогда — то, о чём не знаем, может быть, только по ключевым словам, если повезёт. Правда и это неплохо и удовлетворяет большинство пользователей. Но вот что свойственно западным БД — это практически полное отсутствие русскоязычных работ. Так, в Web Of Science профессор физфака МГУ Постнов К. А. присутствует двумя русскоязычными работами, а академик Черепашук А. М. — одной, притом, что и тот, и другой имеют большой индекс цитирования. Т.е. русскоязычные работы совсем неизвестны западным исследователям. С этими авторами просто — они часто публикуют работы и на английском языке в зарубежных журналах. Для других — это проблема, они неизвестны зарубежному читателю, а многие и русскоязычному, всё по той же причине — мизерные тиражи научных журналов и фактическое отсутствие русскоязычных информационно-поисковых систем (ИПС). У нас в стране хорошими информационно-поисковыми системами можно назвать Научную электронную библиотеку (<http://elibrary.ru/>) и портал MathNet (<http://www.mathnet.ru/>).

Кстати, почти все перечисленные выше ИПС обладают хорошей полнотой архивной информации — западные на десятилетия вглубь, а служба Chemical Abstracts Service и на столетия, наши, по крайней мере, на 2000-е годы. Что ещё в них отсутствует — полноценный рубрикатор, который дает возможность окинуть взглядом проблему целиком. И это можно понять, рубрикация сильно усложняет процесс подготовки документов.

В 1950–1980-е гг. функцию информирования о русскоязычных работах выполнял РЖ ВИНТИ. На него

была хорошая подписка в нашей стране, да и на Западе, в том числе были подписчики не только в соцстранах, но и в США, Великобритании, и, учитывая это, в нем даже одно время публиковались переводы заголовков на английский язык. С 1991 года всё изменилось [3–5] как видно по рис. 1, 2. В то же время, потерявший доступность печатный РЖ ВИНТИ или вовсе прекратившие выходить печатные Chemical Abstracts, Physics Abstracts или Astronomy and Astrophysical Abstracts — небольшая потеря с приходом электронных продуктов, отвечающих изменившимся условиям и новой программно-технической базе. В 2000-х гг. начали появляться архивы научных журналов за все годы их выхода, начавшиеся в нашей стране с «Успехов физических наук» и «Акустического журнал» [6, 7] (мы их отмечаем, т.к. редакции смогли оперативно выложить в Интернет в свободном доступе полнотекстовые архивы всех номеров журналов). Подавляющее число русскоязычных научных журналов в виде библиографий и резюме собрано в удачном проекте РФФИ Научная электронная библиотека (<http://elibrary.ru/>), на основе БД которого составляется и индекс цитирования (РИНЦ) [8, 9].

К нашим двум информационным системам, указанным раньше, конечно, есть некоторые претензии, но они совсем неплохи — есть полнота, есть удовлетворительное качество отражения материала, а в MathNet и просто хорошее, есть оперативность, присутствует архивная часть и, в том числе и поиск по ключевым словам. Нет главного, на наш взгляд — хорошей рубрикации работ. Это даже не претензия. Они не для этого создавались или это не стали делать при их создании, т.к. тогда они, может, и до сих пор бы не появились, и мы не имели бы возможности оперативно знакомиться с новыми работами.

К тому же, нам кажется, по старинке думали, что такую функцию исполняет информационная система «Сокол» ВИНТИ, но:

1. во-первых, попробуйте к ней обратиться. Вот адрес: <http://www2.viniti.ru/>;
2. во-вторых, её БД хоть и датируется 1981 г., но на самом деле по многим тематикам начала наполняться только с 1990-х гг., а по математике — только с 1997 г.;
3. в-третьих, по опыту пользования знаем, что БД, на которой работает эта ИПС, весьма грязна по наполнению, с огромными пропусками по книгам, статьям из журналов (многие журналы представлены далеко не полным содержанием, нередко отсутствуют целые номера), низкой оперативностью отражения материала и ещё с 1990-х гг. совсем слабой квалификацией редакторов по большинству тематических выпусков, несмотря на видимое присутствие «в шапках» фамилий академиков и докторов наук. Бессребренников, работавших ранее за возможность оставлять себе

статьи, не стало, а гонорар «внештатников» не оправдывает затраты даже на дорогу или пересылку рефератов Почтой России.

Отметим также упавшее наполнение РЖ, на основе которого формировалась БД ВИНИТИ, начиная с 1991 г. (например, по РЖ «Физика», см. рис. 3, где приведено реальное наполнение и прогнозируемый по результатам 1954–1990 гг. рост).

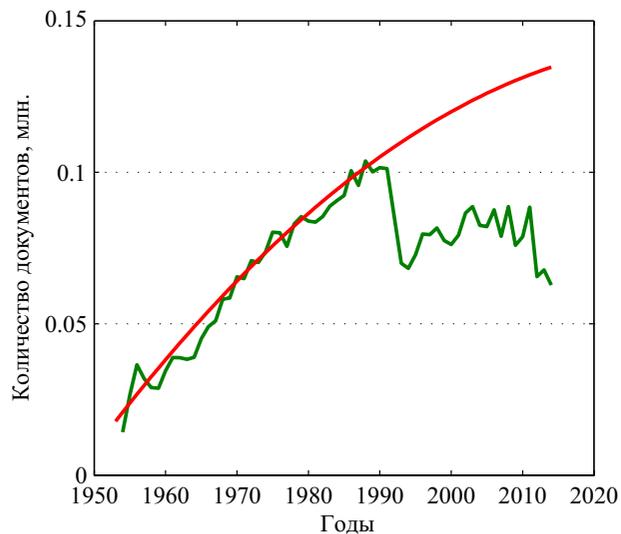


Рис. 3: Наполнение Сводного тома РЖ «Физика», 1954–2012 гг. (синий цвет). Красным цветом обозначена аппроксимация наполнения в 1954–1990 гг. и прогноз наполнения на период с 1991 г.

На рис. 4 приведены также данные по РЖ «Химия» в сравнении с Chemical Abstracts. По рисунку видно, что на рубеже 1955–1956 гг. они были сравнимы по наполнению и сильно разошлись в дальнейшем.

На основе «западных» БД были созданы и пользуются известностью у нас в стране такие информационно-поисковые системы как SciFinder, Web Of Science, База рефератов ADS (NASA), SCOPUS, INSPEC и др. Их БД по зарубежным источникам имеют хорошее наполнение, приличную оперативность, не скажем, что у всех удобный интерфейс, но, главное, они доступны. Например, для сотрудников Московского государственного университета им. М. В. Ломоносова на сайте (<http://www.msu.ru/resources/>), а Санкт-Петербургского государственного университета — <http://cufts.library.spbu.ru/CRDB/SPBGU/browse/facets/name/> (в последнем они довольно глубоко «прикопаны»).

Тупиковая ситуация может сложиться, если нам закроют доступ к этим ресурсам. Маловероятно. Но, например, можно представить себе сценарий помягче — будут резко подняты цены, которые и сейчас нам кажутся чрезмерными, или огрублен доступ к источникам. Как тут быть с прививавшимися нам с ранних

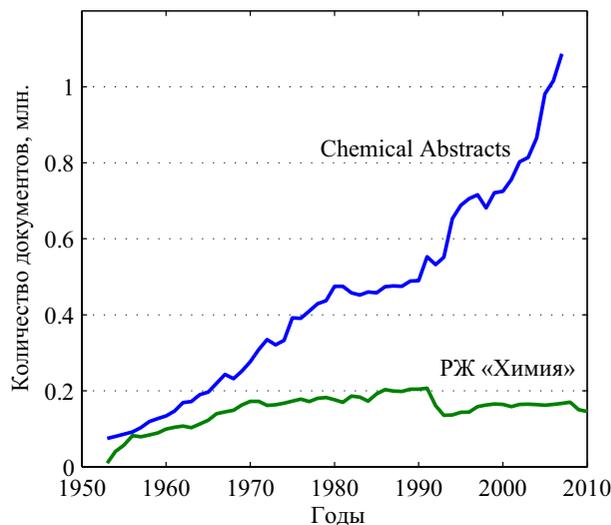


Рис. 4: Наполнение БД Chemical Abstracts и РЖ «Химия» ВИНИТИ

1990-х гг. мантрами, вроде — «наука не имеет границ», «мы все — одна семья», «знания принадлежат всему человечеству» и т. п. Ведь мы же помним, как это было с позиционной системой GPS. Хорошо, среагировали и, худо-бедно, создали ГЛОНАСС. Проблема есть.

А что же у нас сейчас есть. В области информационного обеспечения научных исследований наблюдается неприятная тенденция. Это, прежде всего, начиная с 1990-х гг., резкое снижение тиражей русскоязычных научных журналов, а не их популярности, как мы считаем (рис. 5).

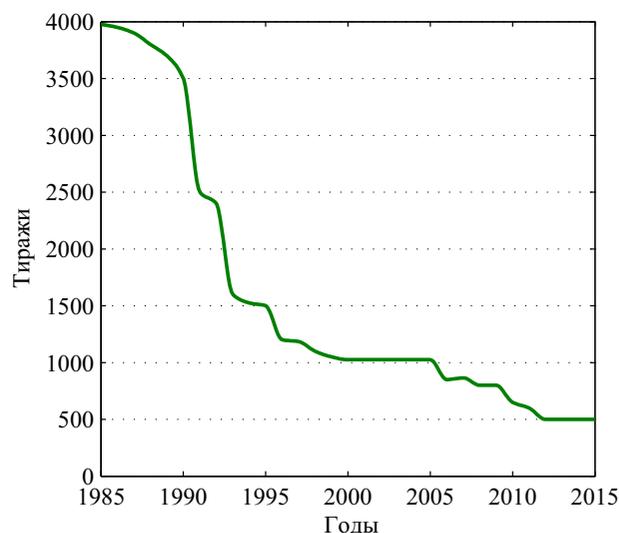


Рис. 5: Тиражи журнала «Успехи физических наук»

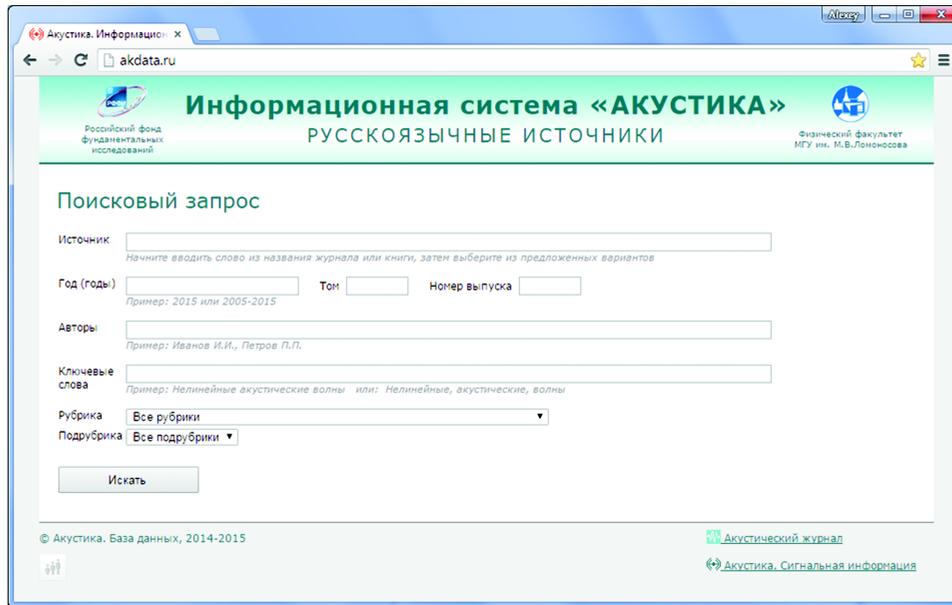


Рис. 6: Информационная система «Акустика». Русскоязычные источники

Это по УФН, а для других научных журналов тиражи в настоящее время колеблются вообще возле сотни экземпляров. Непомерное увеличение цены подписки сделало свое дело. Для УФН, «Акустического журнала» и некоторых других «продвинутых» научных изданий проблема решается наличием электронной версии в Интернете — их сайты являются настоящими информационно-поисковыми системами по этим журналам, но с подавляющей массой журналов проблема доступа остается.

И здесь помогает возникшая в 2000-х гг. идея ИПС, в том числе и тематических.

Основные требования, предъявляемые к ИПС:

1. хорошее наполнение в объявленной научно-технической области (НТО) за указанный временной промежуток;
2. быстрое появление текущей информации;
3. удобный и хорошо продуманный рубрикатор НТО;
4. удобный интерфейс пользователя, отвечающий большинству его запросов, но в то же время не перегруженный малозначительными полями;
5. наличие необходимых, но не навязчивых подсказок и т. д.

Важным качеством таких проектов должно быть постоянное пополнение их 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, снимаются данные для последующего анализа тенденций в развитии той или иной области, производятся другие виды информационного обслуживания, например, выпуск тематической «сигнальной информации».

Вся сложность конструкции и реализации такой ИПС должна лежать вне внимания «потребителей». Для них важны простота, удобство использования и, конечно, получение нужных сведений по запросу, что мы ранее отмечали в [12]. Были попытки поставить между пользователем и ИПС «прокладки» в виде «информационника», который бы по заданию специалиста проводил квалифицированный поиск, но дальше эксперимента дело не пошло. Выяснилось, что лучший результат при информационном поиске получается всё же тогда, когда его проводит не посредник, а сам потребитель информации, если он, конечно, является квалифицированным ученым или инженерно-техническим работником.

2. ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ АКУСТИКИ

Вот на таком фоне авторами начала создаваться ИПС «Акустика» по русскоязычным источникам (<http://akdata.ru/>, рис. 6). Такая языковая специализация вызвана, как уже отмечалось, почти полным отсутствием русскоязычной части в западных ИПС. Для пользователей в Интернете она имеет простой и интуитивно понятный интерфейс.

ИПС «Акустика» на сегодня отражает около 300 периодических изданий, книги и труды 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. Для основных журналов, так называемых полноразметочных, глубина ретроспективы значительно больше.

Проведенное сравнение с поиском в Google Scholar показывает, что по взятым произвольно ключевым словам в начальных страницах его выдачи присутствует только малая часть документов, выдаваемых ИПС «Акустика».

Наряду с ИПС «Акустика» в качестве информационного обеспечения акустических исследований были созданы сайты полнотекстового архива «Акустического журнала» (<http://www.akzh.ru/>) [7] и Сигнальной информации по акустике (<http://akinfo.ru/>) [11]. Все они объединены на портале «Информационная система «Акустика». Русскоязычные источники» (<http://akdata.ru/>).

ЗАКЛЮЧЕНИЕ

Таким образом, на сегодня информационное обеспечение исследований по акустике, созданное нами, включает в себя три проекта:

1. «Информационная система «Акустика». Русскоязычные источники» (<http://akdata.ru/>). Она в настоящее время в значительной степени покрывает 2005–2015 гг., а по «профильным» журналам и все доступные нам годы и содержит около 30 тыс. документов (за 2005–2014 гг. — 20 тыс.). И она реализует все пункты, которые мы отнесли к продвинутому пользователю. Для

сравнения отметим, что за этот период в БД «Акустика» ВИНТИ появилось только около 2000 русскоязычных документов, т. е. на порядок меньше.

2. полнотекстовый «Архив Акустического журнала» в Интернете (<http://www.akzh.ru/>) Поддержка сайта и своевременная актуализация. Подготовка электронных версий текущих выпусков АкЖ: рубрицирование статей, сканирование наполнения журнала, дополнение статьями, не вошедшими в русскоязычную версию, генерирование русскоязычных и англоязычных страниц сайта;
3. проект «Акустика. Сигнальная информация» (<http://akinfo.ru/>). Отражение текущего состояния акустических исследований в русскоязычном мире. Периодичность обновления информации — каждые два месяца. Подготовка электронных версий текущих выпусков Сигнальной информации: поиск и рубрицирование документов, структурирование их библиографии и резюме, ввод подготовленной информации в базу данных, генерирование страниц сайта.

Создание надежных, простых для пользователя и, по возможности, бесплатных информационно-поисковых систем по русскоязычным источникам является на данный момент неотложной задачей, чему, надеемся, и способствуют наши работы на физическом факультете МГУ им. М. В. Ломоносова.

Работа выполнена при финансовой поддержке РФФИ, выделившего грант № 12-07-00732-а.

- [1] Черный А. И. Всероссийский институт научной и технической информации: 50 лет служения науке. (М.: ВИНТИ, 2005). 306 с.
- [2] Шамаев В. Г. Научно-техническая информация. Сер. 1. № 10. С. 21. (2010).
- [3] Семенов В. В. Вестник РАН. **80**, № 4. С. 337. (2010).
- [4] Шамаев В. Г. Вестник РАН. **81**, № 5. С. 430. (2011).
- [5] Шамаев В. Г. Вестник РАН. **83**, № 10. С. 910. (2013).
- [6] Сайт журнала «Успехи физических наук». <http://ufn.ru/>
- [7] Шамаев В. Г., Горшков А. Б., Жаров А. В. Акустический журнал. **59**, № 2. С. 283. (2013).
- [8] Михайлов О. В. Вестник Российской академии наук. **81**,

№ 6. С. 557. (2013).

- [9] Фёдоров П. П., Попов А. И. Вестник Российской академии наук. **84**, № 3. С. 222. (2014).
- [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., Горшков А. Б., Шамаев Н. В. Акустический журнал. **60**, № 1. С. 109. (2014).
- [12] Шамаев В. Г., Шамаев Н. В. Акустический журнал. **57**, № 4. С. 550. (2011).

WEB Information Search Systems and information support for acoustics

V. G. Shamaev^{1,a}, A. B. Gorshkov², L. G. Gushchina³, V. I. Yakimenko⁴

¹Department of Acoustics, Faculty of Physics, M. V. Lomonosov Moscow State University, Moscow 119991, Russia

²Sternberg Astronomical Institute, M. V. Lomonosov Moscow State University, Moscow 119991, Russia

³Kaluga branch of Bauman Moscow State Technical University, Kaluga, 248000, Russia

⁴Tyumen State University, Tyumen, 625003, Russia

E-mail: ^ashamaev08@gmail.com

We address the problem of retrieval of Russian scientific papers in the Web. The absence of these papers in foreign databases leads to their loss to the scientific community. In this connection, attention is drawn to the Russian-language information search system (ISS) — Scientific Electronic Library (<http://elibrary.ru/>), portal MathNet (<http://www.mathnet.ru/>). The Russian-language part of the database VINITI (<http://www2.viniti.ru/>) is criticized. We briefly describe our «Information System «Acoustic». Russian-language sources» (<http://akdata.ru/>). The system contents nearly 30000 papers covering 300 periodicals from country-regionplaceRussia and CIS countries, as well as books and conference proceedings. Particular attention is paid to the full-text archives of «Uspekhi Fizicheskikh Nauk» (<http://ufn.ru/>) and «Akusticheskij Zhurnal» (<http://www.akzh.ru/>), as well as the site of «Current Contents. Acoustics» (<http://akinforu.ru/>).

PACS: 43.10.Pr.

Keywords: Web-resource, Information Search Systems, Current Contents for Acoustics, Russian sources, e-docs, databases.

Received 09.02.2015.

Сведения об авторах

1. Шамаев Владимир Григорьевич — канд. физ.-мат. наук, старший научный сотрудник; тел.: (495) 939-29-21, e-mail: shamaev08@gmail.com.
2. Горшков Алексей Борисович — научный сотрудник; тел.: (495) 939-48-90, e-mail: gorshkov@sai.msu.ru.
3. Гущина Людмила Григорьевна — канд. физ.-мат. наук, доцент.
4. Якименко Владимир Иосифович — старший преподаватель.