

Предпосылки создания комплексированных индексов цитирования сотрудников научного института

© О.Л. Обухова, М.Ю. Заикин, И.В. Соловьев

Институт проблем информатики РАН
obuhova@amsd.com

Аннотация

В статье проводится анализ стандартных библиометрических сервисов, а именно, Thomson Reuters' Web of Science (WoS) и Российского индекса научного цитирования (РИНЦ) на платформе eLibrary, которые используются при оценке профессиональной значимости российских ученых. Особенности этих ресурсов приводят к серьезной недооценке публикационного вклада научных сотрудников. Как одно из возможных решений по преодолению недостатков общепризнанных сервисов в подсчете рейтинга ученых авторы предлагают метод создания комплексированных индексов научного цитирования.

Введение

Термин "наукометрия" [1] для обозначения одного из направлений науковедения, использующего математические методы для изучения процесса развития науки, впервые ввел профессор МГУ В.В. Налимов. В его работах особое внимание уделялось системе библиографических ссылок как особому языку научной информации [2]: «ссылки - это кодовый язык, позволяющий представлять публикации в компактной форме». Изучение структуры этого языка для выявления взаимосвязей и классификации научных работ лежит в основе особого направления наукометрии, которое носит название библиометрия. Статистический анализ публикаций и их цитирования позволяет выявлять закономерности и темпы развития различных отраслей знания и их взаимное влияние, и, в частности, определять рейтинги ученых, отражающие их научную продуктивность.

Именно уровень публикационной активности является одним из показателей результативности

научной организации, алгоритм вычисления которой определен Приказом Министерства образования и науки Российской Федерации от «14» октября 2009 г. № 406 об утверждении:

«Типового положения о комиссии по оценке результативности деятельности научных организаций, выполняющих научно-исследовательские, опытно-конструкторские и технологические работы гражданского назначения, и Типовой методики оценки результативности деятельности научных организаций, выполняющих научно-исследовательские, опытно-конструкторские и технологические работы гражданского назначения» [3].

На основе показателей оценки результативности деятельности научной организации она может быть отнесена к одной из следующих категорий [4]:

- а) 1-я категория - научные организации - лидеры;
- б) 2-я категория - стабильные научные организации, демонстрирующие удовлетворительную результативность;
- в) 3-я категория - научные организации, утратившие научный профиль и перспективы развития.

Составной частью методики оценки научной организации является пункт 1.2 Типовой методики «Публикационная активность» раздела 1 «Научный потенциал и эффективность научных исследований». Для оценки публикационной активности методикой определяется ряд показателей: количество публикаций, цитируемость по каждой публикации и импакт-фактор журнала, где опубликована работа. Учитываются только те журнальные статьи, которые зафиксированы в информационных ресурсах:

- реферативно-библиографическая база данных научного цитирования Web of Science, разработанная Институтом научной информации США,

- Российский индекс цитирования (РИНЦ), функционирующий на платформе www.e-library.ru

Поскольку вопрос оценки деятельности научных институтов поставлен очень серьезно, проведем некоторый анализ предлагаемых баз данных с точки зрения полноты наполнения и точности статистических отчетов.

Труды 13й Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» - RCDL'2011, Воронеж, Россия, 2011.

1. Обзор представительства российских научных изданий в Web of Science

1.1 История создания и структура Web of Science

Большой вклад в становление библиометрии внес американский ученый Юджин Гарфилд [5], прикладным результатом исследований которого явилось создание Института научной информации (Institute for Scientific Information, ISI) в Филадельфии, США, преобразованного в 2008 г. в корпорацию Thomson Reuters.

Данная компания занимается вопросами составления библиографических баз данных научных журнальных публикаций, их индексированием и определением индекса цитируемости, импакт-фактора и других статистических показателей научных работ.

В 1960 г. в ISI была создана база данных по естественнонаучным журналам Science Citation Index (SCI); которая в настоящее время носит название Science Citation Index Expanded (SCIE) и содержит каталоги журнальных публикаций 8060 основных научных изданий по 173 дисциплинам начиная с 1900 г.

В начале 1980-х годов начали развиваться базы данных по общественным наукам Social Sciences Citation Index (SSCI) и гуманитарным дисциплинам Arts & Humanities Citation Index (A&HCI).

В совокупности эти базы данных, содержащие каталоги журнальных публикаций и ссылок на них, именуются Web of Science (WoS) [6]. WoS входит в состав ISI Web of Knowledge [7]. ISI Web of Knowledge обеспечивает доступ к большому ряду баз данных, как входящим в состав Web of Knowledge, так и находящимся за ее рамками, и включает в себя:

- указатели цитированной литературы— *ISI Web of Science*

- указатели текущего содержания публикаций— *ISI Current Contents Connect*

- специализированные указатели— *ISI Chemistry, ISI Proceedings, Derwent Innovations Index, BIOSIS Previews, CAB ABSTRACTS, INSPEC*

- инструменты оценки и анализа научного содержания— *ISI Journal Citation Reports® on the Web, ISI Essential Science Indicators*

- веб-ресурс, находящийся в открытом доступе, содержащий перечни наиболее часто цитируемых ученых и их научных достижений— *HighlyCited.com*.

1.2 WoS : достоинства и недостатки

Масштабно организованные базы данных WoS являются для всего мира основными источниками количественных исследований науки. Но в таком безоговорочно доминирующем позиционировании программных продуктов компании Thomson Reuters на рынке информационной наукометрии «сплетается множество исторических данностей,

предпосылок и следствий, которые на выходе способствуют, скорее, не отражению, а искажению количественных и тем более качественных параметров исследовательской деятельности ученых разных стран и разных специальностей» [8]. Одной из причин неадекватного отображения российской науки в WoS является то, что проиндексированные научные продукты не образуют достаточно представительного подмножества работ российских ученых, и экстраполяция показателей этого подмножества на всю российскую науку приводит к ложным выводам и заблуждениям. Почему же существует проблема создания более или менее полного представительства российской науки в WoS?

1.3 Статус национальных научных изданий в WoS

Первый вопрос, который возникает при использовании данных WoS: в какой мере включенные в нее журналы являются международными, а в какой – национальными. Вопрос о роли «национальной» науки и ее соотношении с «мировой» является достаточно непростым, как с теоретической, так и с эмпирической точки зрения [9]. Многие методические ограничения, способы отбора публикаций, определяемые Thomson Reuters и закладываемые в Web of Science, приводят к тому, что гораздо больше внимания уделяется американским публикациям, и что является безусловным - англоязычным работам. В то время, как в таких странах, как Россия, Япония, Китай и т.д., т.е. в странах с развитой экономикой, научное сообщество использует для обмена научными идеями родной язык. Национальные ресурсы научных публикаций формируются на государственном языке страны. Тем не менее, большое количество национальных изданий выходит на английском языке. Проанализируем, как распределяется представительство научных журналов разных стран в WoS.

Если выполнить выборочно запрос к Master Journal List [10], одному из аналитических сервисов ISI Web of Knowledge, с целью получения списка научных журналов, реферируемых в WoS и издающихся в разных странах, то получим результат, представленный на Рисунке 1.

Доля российских журналов составляет около 2%. Почти половину всех включенных в указатель журналов составляют американские издания, точнее, доля изданий США, Англии и Нидерландов равна 61%.

Список отечественных журналов, включаемых в Journal Citation Reports 2009, состоит из 168 наименований, хотя только рекомендованных ВАК для опубликования результатов докторских диссертаций отечественных научных журналов насчитывается более 2 тысяч, причем, следует

Journal Citation Reports®

Journal Summary List

Journals from: countries/territories UNITED STATES

Journals 2697

Journals from: countries/territories ENGLAND

Journals 1558

Journals from: countries/territories NETHERLANDS

Journals 653

Journals from: countries/territories GERMANY

Journals 544

Journals from: countries/territories JAPAN

Journals 206

Journals from: countries/territories FRANCE

Journals 187

Journals from: countries/territories SWITZERLAND

Journals 168

Journals from: countries/territories RUSSIA

Journals 147

Journals from: countries/territories INDIA

Journals 93

Journals from: countries/territories BRAZIL

Journals 89

Рис. 1

отметить, что эти издания проходят рецензирование и поэтому отвечают всем необходимым требованиям к уровню научных статей.

Большинство российских изданий предоставляет переводы содержания своих научных сборников на английский язык. Но это не означает их индексирования в WoS, вследствие чего наблюдается недостаточное присутствие в БД SCIE российских работ, как, собственно, и работ из других стран мира.

1.2 Представительство российских научных журналов, посвященных математике и информатике

Издания по математике и информатике представлены в WoS совсем незначительно. Распределение российских журналов по тематикам в БД Scopus и WoS были проанализированы в работе [11].

Таблица 1. Распределение российских журналов по тематикам в БД Scopus и WoS.

Тематика	Scopus (%)	WoS / SCIE (%)
Биология +	24,5	31,5
Медицина		
Физика	16,0	14,8
Химия	13,3	13,3
Техника (+)	26,2	13,6

Металлургия)		
Науки о Земле	9,5	11,0
Математика + Компьютеры	7,8	8,6
Прочие	2,7	7,5

На долю российских публикаций по математике и информатике в WoS приходится только 8,6% от общего числа отечественных научных работ. Тогда как в нашей стране из более 5700 российских научных журналов, размещенных на платформе eLIBRARY.RU, представляющей Российский индекс научного цитирования, рубрика «Математика» представлена 908 журналами, «Информатика» - 306, «Кибернетика» - 655, что составляет 30% от общего числа изданий. К этому списку можно добавить и рубрику «Автоматика и вычислительная техника», представленную 984 журналами. И только 14 российских журналов, в которых публикуются работы, связанные с тематиками математика и информатика, представлены в WoS.

Таблица 2. Ведущие журналы по математике и информатике по трехлетнему импакт-фактору.

Название издания	Импакт-фактор
Информатика и ее применения	0,825
Искусственный интеллект и принятие решений	0,808
Российские нанотехнологии	0,779
Вопросы когнитивной лингвистики	0,601
Дискретный анализ и исследование операций Институт математики им. С.Л. Соболева СО РАН	0,545
Проблемы теории и практики управления	0,502
Вычислительные технологии	0,501
Известия Российской академии наук. Теория и системы управления (WoS) JOURNAL OF COMPUTER AND SYSTEMS SCIENCES INTERNATIONAL	0,467
Журнал вычислительной математики и математической физики	0,434
Доклады Академии наук, серия Математика. (WoS) DOKLADY MATHEMATICS	0,388
Доклады Академии наук. Серия Науки о Земле. (WoS) DOKLADY EARTH SCIENCES	0,388
Проблемы анализа риска	0,387
Проблемы управления	0,367
Сибирский математический журнал. Институт математики им. С.Л. Соболева СО РАН. (WoS) SIBERIAN MATHEMATICAL JOURNAL	0,359
Математические труды. Институт математики им. С.Л. Соболева СО РАН	0,353
Экономика и математические методы	0,343
Прикладная математика и механика	0,309
Автоматика и телемеханика. (WoS)	0,306

AUTOMATION AND REMOTE CONTROL	
Труды института математики и механики УрО РАН	0,301
Радиотехника и электроника. (WoS) JOURNAL OF COMMUNICATIONS TECHNOLOGY AND ELECTRONICS	0,296
Теоретическая и математическая физика	0,291
Информационные технологии и вычислительные системы	0,282
Вычислительные технологии Институт вычислительных технологий Сибирского отделения РАН	0,244
Электросвязь	0,245
Вычислительные методы и программирование: новые вычислительные технологии	0,232
Сибирский журнал вычислительной математики	0,224
Алгебра и логика	0,223
Информационно-управляющие системы	0,215
Геоинформатика / Geoinformatika	0,211
Информационные ресурсы России	0,210
Программирование. (WoS) PROGRAMMING AND COMPUTER SOFTWARE	0,204
Известия РАН. Серия математическая. (WoS) IZVESTIYA MATHEMATICS	нет
Математические заметки. Ежемесячный журнал Отделения математических наук РАН. (WoS) MATHEMATICAL NOTES	нет
Издается только на англ. языке. (WoS) RUSSIAN JOURNAL OF NUMERICAL ANALYSIS AND MATHEMATICAL MODELLING	нет
Успехи математических наук. (WoS) RUSSIAN MATHEMATICAL SURVEYS	нет
Теория вероятностей и ее применения. (WoS) THEORY OF PROBABILITY AND ITS APPLICATIONS	нет
Вестник РАН. (WoS) HERALD OF THE RAS	нет

Символами WoS отмечены те издания, которые индексируются БД Web of Science

Таблица 2 наглядно демонстрирует тот факт, что наиболее авторитетные и имеющие наибольший импакт-фактор российские издания по математике и информатике лишь в малой мере представлены в WoS. Есть, безусловно, объективные причины незначительного представительства российских научных изданий России в SCIE. Много авторитетных журналов не переводятся на английский язык. В частности, перечень переводимых журналов был сформирован в советское время, когда бразды правления были в руках государственного издания «Наука». В последние же годы появилось много новых журналов, в основном по информатике, которые так и не смогли «пробиться» на западный рынок,

поскольку издаются в разных, не слишком мощных издательствах, которые не в состоянии достичь соответствующих соглашений с западными партнёрами. Но общий вывод очевиден: российская наука не представлена должным образом на WoS.

1.3 Игнорирование WoS публикаций по информатике как глобальная проблема

«Несправедливое» отношение к индексированию работ по информатике серьезно волнует сообщество ученых всего мира в области компьютерных технологий. В статье [13] авторы утверждают, что если для оценки эффективности труда научного работника будут использоваться стандартные библиографические сервисы, а именно, Thomson Reuters' Web of Science (WoS) и Scopus, то большинство работ исследователей в области компьютерных наук не будут учтены при оценке профессиональной значимости ученых. Проведя исследования по соотношению реального списка публикаций ученых из ведущих университетов США и того перечня, который представлен в WoS, авторы доказали, что 66% публикаций по информатике не учитываются в WoS. Причем, этот показатель гораздо выше, чем для ученых в области математики или физики.

К явным недостаткам WoS авторы относят неполноценный охват изданий. До 2008 г. WoS вообще не включало в свою БД труды конференций, которые не были опубликованных в Springer's *Lecture Notes in Computer Science* series. С середины 2009 г. Thompson Reuters включили новый сервис, позволяющий индексировать около 12000 трудов конференций. Но критерии отбора не позволяют говорить о достаточной научной значимости этой подборки. Игнорируются важные издания, к примеру, издаваемые Communication ACM [14].

Поскольку другие библиографические сервисы, такие как DBLP [15], ACM Digital Library [16], CiteSeerX [17], ориентированные на индексацию и учет цитирования публикаций по информатике, не предназначены для сравнения с работами ученых из других областей знаний, то они не выступают в роли общепринятых систем индексации. Также как и Google Scholar [18]. Они «непрозрачны» в отношении того, как в них собирается и обрабатывается библиографическая информация. Вывод единственный: понимая, что WoS и Scopus [19] единственные узаконенные системы индексирования научных трудов, необходимо ученым объединиться в своих усилиях воздействия на WoS в направлении исправления недочетов, недостатков и искажений статистики тех продуктов, которые предоставляются компанией Thomson Reuters.

2. Достоинства и недостатки РИНЦ

Вопросы индексирования публикаций российских ученых и учет их цитирования в достаточно представительном объеме мог бы

решить российский ресурс, реализующий задачи Web of Science в отечественной среде. Большие надежды возлагались на Российский индекс научного цитирования РИНЦ, база данных которого представлена на web-ресурсе eLibrary.ru [20]. Но находящаяся в процессе становления структура Российского индекса научного цитирования подвергается серьезной критике. Директор Библиотеки по естественным наукам РАН профессор Н.Е. Каленов направил письмо [21] в конце прошлого года министру образования и науки А.А. Фурсенко, в котором доказательно показывает, что имеются принципиальные ошибки в самой алгоритмической основе базы данных eLibrary.ru. Данный фактор существенным образом сказывается при подсчете цитируемости и искажает результаты, полученные при различных способах поиска публикации. Работы цитируются как разные, если есть разночтения в библиографических описаниях статьи в пристатейных списках литературы.

2.1 Неполнота информационного наполнения БД

Количество публикаций сотрудников ИПИ РАН за 2010 г. в изданиях, индексируемых РИНЦ, составляет больше 300 работ (включая патенты на изобретения). В БД eLibrary.ru на май месяц содержалась информация только о 68 работах.. Т.е. проиндексировано только 23% работ.

Есть объективные причины того, что eLibrary.ru взятую на себя задачу не выполняет в полном объеме в настоящее время. Издательства передают каталоги журналов с запозданием. Часть ошибок возникает при сканировании исходных материалов, предоставляемых издательствами, в частности, из-за особенностей форматирования pdf- файлов, к примеру, обязательных знаков переносов и «пропадающих» пробелах в пристатейных списках литературы. Как следствие, названия публикаций, выделенные из списков ссылок на цитируемые работы, переносятся в БД в искаженном виде. Очевидно, что данное обстоятельство препятствует созданию полной и достоверной информации как в отношении списков публикаций автора, так и, главное, в отношении формирования правильной системы ссылок на опубликованные и проиндексированные статьи.

2.2 Принципиальные ошибки в алгоритмах составления индексов цитирования

За созданием БД eLibrary.ru стоит, безусловно, большая работа. Но результаты выбора определенной информации из БД часто бывают противоречивы. Возможно, чтобы избежать серьезных нареканий, разработчики пошли по пути предоставления авторам возможности самим формировать свои списки публикаций, введя понятия «привязанных» - «непривязанных» публикаций и ссылок.

Администрация eLibrary предлагают авторам самим просматривать и корректировать списки своих публикаций и цитирований. К примеру, если автор знает, что статья его, но она почему-то не привязана, он сам может привязать эту статью, или ссылку на нее» (из руководства пользователей eLibrary).

Но здесь возникает ряд вопросов:

- 1) у авторов нет возможности скорректировать ошибочные данные о названиях, соавторах, выходных данных его публикаций в БД. Сообщения об этих ошибках предлагается отправлять модераторам. Но смогут ли и будут ли модераторы разбираться в каждом случае? Пока практика показывает, что письма, направленные на <mailto:support@elibrary.ru>, часто остаются без ответа.
- 2) если какой-то зарегистрированный автор по ему известным причинам прикрепил к своему перечню работы из списка «непривязанных» публикаций работы однофамильцев, то как это будет отслеживаться?
- 3) если работы, заведомо принадлежащие автору, уже прикреплены к автору-однофамильцу, то как разбираться в этой ситуации?
- 4) если хотя бы 10% из 588042 зарегистрированных авторов одновременно обратятся БД eLibrary, смогут ли поддерживать столько сессий на редактирование БД?
- 5) Если списки «привязанных» публикаций автор может формировать, то списки ссылок на его работы в соответствии с алгоритмами работы БД ресурса eLibrary пользовательской обработке не подлежат.

На рисунке 2 представлен коллаж из screenshots страниц eLibrary.ru, отражающий списки публикаций автора с редкой фамилией Козмидиади В.А., сделанных с интервалом в один месяц. Данный пример наглядно демонстрирует, что алгоритмы, заложенные в работу eLibrary, по которым происходит сопоставление автор – публикация, не работают корректно. Статьи автора попадают то в «привязанные», то в «непривязанные» публикации и с разным числом цитирований.

Понять критерии такой «перетасовки» не представляется возможным. Но надо отдать должное разработчикам этой системы: общее количество статей в списках «привязанных» и «непривязанных» публикаций остается постоянным 8/4. Но откуда берутся эти константы? Статистические отчеты ресурс eLibrary формирует только по «привязанным» публикациям и ссылкам, следовательно, данный ресурс предоставляет недостоверную информацию. Но, в таком случае, утрачивается основной смысл данного ресурса, который позиционируется как РИНЦ.

КОЗМИДИАДИ В А - СПИСОК ПУБЛИКАЦИЙ

Автор: **КОЗМИДИАДИ ВЛАДИМИР АЛЕКСАНДРОВИЧ**
Институт проблем информатики РАН, отдел 2В (Москва)

Показывать: только включенные в список работ автора (привязанные) публикации
с учетом публикаций, извлеченных из списков литературы

Сортировка: по дате выпуска | Порядок: по убыванию | Очистить | Поиск

Всего найдено публикаций: 8. Показано на данной странице: с 1 по 8

№	Публикация	Цит.
1	РЕЗЕРВНОЕ КОПИРОВАНИЕ, ИСПОЛЬЗУЮЩЕЕ СНИМКИ Козмидади В.А. Информатика и ее приложения. 2005. Т. 3. № 2. С. 15-24.	0
2	СЕРВИСНЫЕ ПРИЛОЖЕНИЯ ДЛЯ РАБОТЫ С ИЕРАРХИЧЕСКИМИ ДАННЫМИ В РАСПРЕДЕЛЕННОЙ ФАЙЛОВОЙ СИСТЕМЕ Дьячков В.А., Захаров В.Н., Козмидади В.А., Кузнецов А.В., Попов А.С., Шугатников Д.С. Системы и средства информатики. 2003. Т. 15. № 1. С. 20-35.	0
3	ПЛАНИРОВАНИЕ ВЫПОЛНЕНИЯ ЗАДАНИЙ СЕРВИСНЫХ ПРИЛОЖЕНИЙ В РАСПРЕДЕЛЕННОЙ СРЕДЕ Захаров В.Н., Козмидади В.А., Кузнецов А.В., Попов А.С., Шугатников Д.С. Системы и средства информатики. 2003. Т. 15. № 1. С. 36-48.	0
4	АЛГОРИТМЫ ПАРАЛЛЕЛЬНОГО ВЫПОЛНЕНИЯ ЗАДАНИЙ СЕРВИСНЫХ ПРИЛОЖЕНИЙ В РАСПРЕДЕЛЕННОЙ СРЕДЕ Дьячков В.А., Захаров В.Н., Козмидади В.А., Кузнецов А.В., Шугатников Д.С. Системы и средства информатики. 2005. Т. 15. № 1. С. 49-70.	0
5	УПРАВЛЕНИЕ ВЫПОЛНЕНИЕМ ЗАДАНИЙ СЕРВИСНЫХ ПРИЛОЖЕНИЙ В РАСПРЕДЕЛЕННОЙ СРЕДЕ НА БАЗЕ РЕСУРСНОЙ МОДЕЛИ Дьячков В.А., Захаров В.Н., Козмидади В.А., Кузнецов А.В., Попов А.С. Системы и средства информатики. 2005. Т. 15. № 1. С. 71-81.	0
6	СРЕДСТВА ОБЕСПЕЧЕНИЯ ОТКАЗУСТОЙЧИВОСТИ ПРИЛОЖЕНИЙ Захаров В.Н., Козмидади В.А. Информатика и ее приложения. 2007. Т. 1. № 1. С. 14-25.	1
7	РЕАЛИЗАЦИЯ ОТКАЗУСТОЙЧИВОСТИ СЕРВЕРОВ ПРИЛОЖЕНИЙ Захаров В.Н., Козмидади В.А. Наукосфера: новые технологии. 2006. Т. 7. № 5. С. 56-62.	1
8	ОПИСАНИЕ НЕДЕТЕРМИНИРОВАННОГО ПОВЕДЕНИЯ И ПРОБЛЕМА ОТКАЗУСТОЙЧИВОСТИ ПРИЛОЖЕНИЙ Захаров В.Н., Козмидади В.А., Кузнецов А.В.	0

КОЗМИДИАДИ В А - СПИСОК ПУБЛИКАЦИЙ

Автор: **КОЗМИДИАДИ ВЛАДИМИР АЛЕКСАНДРОВИЧ**
Институт проблем информатики РАН, отдел 2В (Москва)

Показывать: только включенные в список работ автора (привязанные) публикации
с учетом публикаций, извлеченных из списков литературы

Сортировка: по дате выпуска | Порядок: по убыванию | Очистить | Поиск

Всего найдено публикаций: 8. Показано на данной странице: с 1 по 8

№	Публикация	Цит.
1	РЕЗЕРВНОЕ КОПИРОВАНИЕ, ИСПОЛЬЗУЮЩЕЕ СНИМКИ Козмидади В.А. Информатика и ее приложения. 2005. Т. 3. № 2. С. 15-24.	0
2	ШИРОКОВЕЩАТЕЛЬНАЯ РАССЫЛКА СООБЩЕНИЙ, ОПИРАЮЩАЯСЯ НА ПРИМИТИВЫ РАСПРЕДЕЛЕННОЙ ФАЙЛОВОЙ СИСТЕМЫ Козмидади В.А. Системы и средства информатики. 2009. Т. 19. № 1. С. 63-74.	0
3	РАБОТА С КООРДИНАЦИОННОЙ ИНФОРМАЦИЕЙ В РАСПРЕДЕЛЕННЫХ СИСТЕМАХ Козмидади В.А. Системы и средства информатики. 2009. Т. 19. № 1. С. 75-86.	0
4	СРЕДСТВА ОБЕСПЕЧЕНИЯ ОТКАЗУСТОЙЧИВОСТИ ПРИЛОЖЕНИЙ Захаров В.Н., Козмидади В.А. Информатика и ее приложения. 2007. Т. 1. № 1. С. 14-25.	1
5	РЕАЛИЗАЦИЯ ОТКАЗУСТОЙЧИВОСТИ СЕРВЕРОВ ПРИЛОЖЕНИЙ Захаров В.Н., Козмидади В.А. Наукосфера: новые технологии. 2006. Т. 7. № 6. С. 55-62.	1
6	РЕАЛИЗАЦИЯ ОТКАЗУСТОЙЧИВОСТИ СЕРВЕРОВ ПРИЛОЖЕНИЙ. Захаров В.Н., Козмидади В.А. Системы высокой доступности. 2006. Т. 2. С. 56	2
7	МЕТОДЫ РЕАЛИЗАЦИИ ОТКАЗУСТОЙЧИВОСТИ ПРИЛОЖЕНИЙ С НЕДЕТЕРМИНИРОВАННЫМ ПОВЕДЕНИЕМ Долгополов В.С., Захаров В.Н., Козлова И.М., Козмидади В.А., Обухова С.П. Системы и средства информатики. 2006. Т. 16. № 1. С. 374-385.	0
8	ОПИСАНИЕ НЕДЕТЕРМИНИРОВАННОГО ПОВЕДЕНИЯ И ПРОБЛЕМА ОТКАЗУСТОЙЧИВОСТИ ПРИЛОЖЕНИЙ Захаров В.Н., Козмидади В.А., Кузнецов А.В.	0

Рис.2

3. В поисках выхода

При сложившейся на сегодняшний день ситуации с получением информации о публикационной активности и цитировании работ научных сотрудников, авторы статьи видят необходимость создания альтернативной программной системы, целью которой является формирование комплексированных индексов цитирования сотрудников научного института. Для создания Интернет-ресурса, с одной стороны,

функционально полного, а с другой, – без правовых нарушений, необходимо решить, по крайней мере, две достаточно серьезные проблемы – технологическую и организационную. Технологическая проблема сравнительно проста и заключается в выборе программных и технических средств реализации БД. Источниками информации для наполнения базы данных альтернативной системы являются научные публикации сотрудников научного института. Можно опираться на две достоверные составляющие:

- точный список сотрудников

- полный перечень работ, включающий не только журнальные публикации, но и монографии, книги, сборники трудов конференций и книжные серии, учебники, патенты, зарегистрированные программные продукты.

Для формирования системы цитирований предполагается разработка программ-роботов, которые будут использовать web-интерфейс для формирования запросов к БД eLibrary.ru и WoS, реализованный на HTML – формах внешних библиотек. Полученные страницы результатов будут обрабатываться специальными программами-parser, что позволит устранить синтаксические ошибки в списках ссылок на цитируемые работы, с чем пока не справляются функционалы eLibrary.ru. Отредактированные записи будут пополнять БД. Предлагаемый подход использования web-интерфейса для формирования запросов к внешним библиотекам позволит подключать разнообразные тематические ресурсы для процесса наполнения БД. Но надо понимать, что достоверность информации из этих источников не является 100%.

3.1 Пользовательский интерфейс

Пользовательский интерфейс целесообразно организовать на принципах фасетной навигации [22]. В предлагаемом подходе для каждой научной публикации при занесении в коллекцию создается описание её семантической структуры в форме *фасетной формулы объекта*, представленной в виде множества совокупностей: <фасетный признак : список значений> для данного объекта. Научная публикация совместно с *фасетной формулой объекта* называется информационным объектом (ИО) [23]; объединение всех фасетных формул определяет семантическое содержание коллекции научных публикаций.

3.2 Программная среда реализации

Общедоступные и распространенные программные средства, которые авторы предполагают использовать в своей реализационной модели, включают в себя следующие компоненты:

- реализация осуществляется в среде ОС: Microsoft Windows 2000/2003;
- веб-сервер: Microsoft Internet Information Server 5.0/6.0;
- контейнер серверного приложения: Microsoft ASP.NET 2.0;
- СУБД: Microsoft SQL Server 2005;
- в качестве средств разработки используется Microsoft Visual Studio 2005, ASP.NET 2.0;
- для построения пользовательского интерфейса используется Adobe Flash со встроенным языком ActionScript;

- для построения фасетных формул пользовательским интерфейсом используется ASP.NET Web Services [24]

ASP.NET Web Services используют индустриальные стандарты:

- XML. Обмен данными между пользовательским интерфейсом и сервером производится в формате XML [25]

- SOAP. Протокол обмена сообщениями между Web службой и клиентом, основанный на XML [26].

- Web Services Description Language (WSDL). Описывает параметры сообщений Web-службы для взаимодействия с клиентами [27].

Опытная эксплуатация реализационной модели позволит вести файлы протокола и собирать статистическую информацию по подсчету индексов цитирования и сопоставления их с соответствующими показателями из WoS и eLibrary.

3.3 Административный режим

Наполнение и редактирование БД производится в административном режиме администраторами системы и зарегистрированными пользователями-модераторами. Разделение прав доступа позволит отслеживать адекватное наполнение БД.

В административном режиме формируется и редактируется база данных информационных объектов. Поддерживаются следующие функции:

- занесение нового информационного объекта,
- изменение свойств информационного объекта,
- удаление информационного объекта.

Вместе с размещением публикации пользователь заполняет анкету, значения полей которой определяют значения соответствующих фасетных признаков. Авторы планируют следующий сценарий заполнения анкеты.

• Для каждого поля предлагается список возможных значений, в котором пользователь выбирает произвольное количество значений, которые, на взгляд пользователя, отражают смысловое содержание занесенного ИО. Иными словами, пользователь в формализованном и структурированном виде готовит аннотацию своего документа.

• Пользователь может послать запрос Главному Администратору системы о расширении списка возможных значений. После утверждения экспертами данное значение включается в общий список в административном режиме.

Введение нового значения приводит к обновлению визуального интерфейса поисковой системы.

3.4 Создание комплексированных индексов научного цитирования

Создание альтернативной программной системы индексации публикаций сотрудников ИПИ РАН позволит обеспечить полноту наполнения базы

данных с точки зрения учета всех авторов и полного списка публикаций, включающего не только журнальные статьи, но и монографии, учебники, книги, патенты. Система администрирования предоставляет возможность постоянной актуализации списка работ и ссылок на них, используя ресурсы не только баз индексирования WoS и eLibrary, но и возможности поисковых систем Google, Yandex. и т.д., но, правда, с учетом их неполной достоверности, требующей дополнительной проверки.

При использовании технологии фасетной навигации реализуется возможность визуального сопоставления индексов научного цитирования на базе показателей WoS, eLibrary и альтернативной программной системы. Комплексованных индекс может быть построен с учетом каждого из этих трех показателей.

Альтернативная программная система, с одной стороны, обеспечит сохранность и полное отражение научных трудов Института в единой базе данных, с другой стороны, предоставит необходимые сведения для более точной оценки публикационной активности каждого научного сотрудника для сверки данных, формируемых в отчетах РИНЦ. Дополнительно можно реализовать сервисы, целью которых является оптимизация информационного обслуживания сотрудников. Для достижения указанных целей целесообразно строить архив на технологической основе архивов открытого доступа, создавая описания документов в формате, обеспечивающем поддержку многообразных связей между документами и их частями.

Для обеспечения правовой чистоты БД служит поддержка механизмов дифференциации доступа. Бразды правления остаются в руках Главного Администратора, задачей которого является предотвращения появления неточной, или недостоверной, или не подкрепленной реальными документами информации в БД.

Заключение

Поскольку альтернативная программная системы не имеет правового статуса, то использование её показателей возможно только для узконаправленных целей достижения объективности в оценке эффективности труда интеллектуальных работников.

Хотя и объективность носит все-таки относительный характер. Есть целый ряд непреодолимых факторов, исподволь влияющих на общие показатели цитирования научных работ. К примеру, принадлежность к научному «мэйнстриму». Или активность научного работника, включающая возможность большого количества выступлений на конференциях и потока публикаций, его контакты, напористость, дружеские связи, удача.

Но если возможна полемика с Министерством образования и науки в отношении того, какие сервисы, индексирующие научные публикации, должны носить законодательный характер, то альтернативная система может предоставлять документально подтвержденные корректировки статистических отчетов, формируемых WoS и РИНЦ, реализованной на платформе eLibrary.

Но пока у научных сотрудников есть только один из законных способов отстаивать свои права и позиции в научном мире и системах индексирования - это внимательно относиться к работам коллег и возможности их цитирования, и, главное, грамотно и аккуратно составлять списки ссылок на цитируемую литературу.

Литература

- [1] Налимов В.В., Мульченко З.М. Наукометрия: Изучение развития науки как информационного процесса. М.: Наука, 1969. 192 с
- [2] Налимов В.В., Мульченко З.М. Об использовании статистических методов при управлении развитием науки //Управление, планирование и организация научных и технических исследований. Т. 3. М.: ВИНТИ, 1970. С. 327-342
- [3] Приказ Министерства образования и науки Российской Федерации от « 14 » октября 2009 г. N 406 «Об утверждении типового положения о комиссии по оценке Результативности деятельности научных организаций, выполняющих научно-исследовательские, опытно-конструкторские и технологические работы гражданского назначения , и типовой методики оценки результативности деятельности научных организаций, выполняющих научно-исследовательские, опытно конструкторские технологические работы гражданского назначения». Зарегистрирован в Минюст России от 28 января 2010 г. N 16115. - <http://mon.gov.ru/>
- [4] Постановлением Правительства Российской Федерации от 8 апреля 2009 г. N312 «Правила оценки результативности деятельности научных организаций, выполняющих научно-исследовательские, опытно-конструкторские и технологические работы гражданского назначения». - <http://www.government.ru/gov/results/7006/>
- [5] Eugene Garfield. - <http://www.garfield.library.upenn.edu/>)
- [6] Web of Science. - http://thomsonreuters.com/products_services/science/science_products/a-z/web_of_science
- [7] ISI Web of Knowledge - <http://isiwebofknowledge.com/>
- [8] Митрошилова Н.В. Недоброкачественные сегменты наукометрии // Вестник Российской Академии наук. 2011. Том 81, № 2. С. 134–146.

- [9] Савельева И.М., Полетаев А.А. Публикации российских авторов в зарубежных журналах по общественным и гуманитарным дисциплинам в 1993–2008 гг.: количественные показатели и качественные характеристики // Препринт WP6/2009/02 . — М.: Изд. дом Государственного университета – Высшей школы экономики, 2009. — 52 с.
- [10] Master Journal List // <http://admin-apps.isiknowledge.com/JCR/JCR>
- [11] Н.С. Солошенко, О.В. Кириллова. Отражение российских журналов в БД Science Citation Index и SCOPUS // *Educational Technology & Society* 9(3) 2006 ISSN 1436-4522
- [12] Общероссийский математический портал Math-Net.Ru // <http://www.mathnet.ru/>
- [13] Jacques Wainer, Siome Goldenstein, Cleo Billa. Invisible Work in Standard Bibliometric Evaluation of Computer Science // *Communications of the acm*, May 2011, vol. 54, no. 5
- [14] Association for Computing Machinery. - <http://www.acm.org/>
- [15] DBLP - <http://www.informatik.uni-trier.de/~ley/db/>
- [16] ACM Digital Library <http://portal.acm.org/dl.cfm>
- [17] CiteSeerX - <http://citeseerx.ist.psu.edu/>
- [18] Google Scholar - <http://scholar.google.com/>
- [19] Scopus. Библиографическая и реферативная база данных и инструмент для отслеживания цитируемости статей, опубликованных в научных изданиях - <http://www.info.sciverse.com/scopus/>
- [20] РИНЦ // www.elibrary.ru
- [21] Каленов Н.Е. Ещё раз о РИНЦ. - <http://trv-science.ru/2011/02/01/eshhe-raz-o-rinc/#more-10080>
- [22] Обухова О.Л., Соловьев И.В., Бирюкова Т.К., Гершкович М.М., Чочиа А.П. Модель фасетного информационного поиска в коллекции научных материалов. // "Системы и средства информатики", доп. выпуск, , Москва, Наука, 2009 стр. 163 – 174
- [23] Обухова О.Л., Бирюкова Т.К., Гершкович М.М., Соловьев И.В., Чочиа А.П. «Метод динамического создания связей между информационными объектами базы знаний» // Труды 11-ой Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» - RCDL'2009, Петрозаводск, Россия, 2009, стр. 39 –45
- [24] XML Web Services Basics - <http://msdn.microsoft.com/enus/library/ms996507.aspx>
- [25] XML - <http://www.w3.org/XML/>
- [26] SOAP Version 1.2 - <http://www.w3.org/TR/soap/>
- [27] Web Services Description Language (WSDL) Version 2.0 - <http://www.w3.org/TR/wsdl20/>

The Preconditions of a New Method of Complexed Science Citation Indexes for Researchers of Scientific Institute

© Olga Obuhova, Maxim Zaikin, Ivan Soloviev

The content of this article is the analysis of common bibliometric services like Thomson Reuters' Web of Science (WoS) and the Russian scientific Citation Index (RSCI) based on eLibrary platform, which are used for evaluation of Russian scientists professional rating.

The specificity of these services often leads to considerable undervaluation of the Russian scientist publication contribution.

The authors offer the method of complex indexes of scientific quotation as one of the ways to remove the faults of known services of the scientific ratings calculation.