

# Влияние библиометрических методов на формирование рейтинга научной организации

© В.Н. Гуреев

Государственный научный центр вирусологии и биотехнологии «Вектор», Новосибирская область, Кольцово

gureyev@vector.nsc.ru

© Н.А. Мазов

Институт нефтегазовой геологии и геофизики им. академика А.А. Трофимука СО РАН, Новосибирск

MazovNA@ipgg.sbras.ru

## Аннотация

В статье обсуждается возможная роль научных библиотек в формировании рейтинга научных организаций и университетов, отраженного в международных библиометрических базах данных, таких как Web of Science и Scopus. Авторы выделили три комплекса проблем, решение которых могло бы способствовать лучшей представленности информации о российских организациях в международных базах данных. Во-первых, это создание внутренних реферативных баз данных о публикациях сотрудников организации, которые открывали бы научному сообществу доступ к информации о работе учреждения. Во-вторых, это непосредственное редактирование профилей научных организаций в международных библиометрических базах данных. В-третьих, это разработка новых методов информационного обеспечения научных сотрудников актуальной информацией, что опосредованно может стимулировать создание новых качественных публикаций и тем самым повышать рейтинг научной организации. В статье приводится конкретный опыт применения описанных методик на базе двух научных институтов.

Рейтингу научных и образовательных организаций, который в частности выражен количественными показателями публикационной активности сотрудников организаций, зафиксированными в библиометрических базах данных, в настоящее время уделяется повышенное внимание. Наукометрические базы данных, и мировые, и национальные, за исключением сайтов организаций, являются практически единственным общедоступным источником информации о качестве работы учреждения, его рейтинге.

В силу различных причин рейтинги российских научных учреждений и университетов на общемировом фоне являются относительно низкими. Эта ситуация обсуждалась на высшем уровне руководства страны и вылилась в подготовку постановлений, направленных на улучшение представленности российской науки в мировых базах данных [13, 12]. Основные решения, которые могут оказаться эффективными и в результате привести к перелому ситуации, лежат в области повышения финансирования науки и образования в целом, отдельных ученых, в создании переводческих сервисов при организациях для продвижения статей в зарубежные журналы, в продвижении самих российских журналов на западный рынок научной информации.

Вместе с тем, на наш взгляд, можно выделить три комплекса проблем, находящихся в сфере компетенции научных библиотек, решение которых может также оказать существенное воздействие на улучшение рейтинга научных и образовательных организаций, поскольку именно в научных библиотеках сосредоточена работа с научными публикациями и библиометрическими вычислениями.

## 1. Институциональные базы данных

Низкую представленность российских статей в наукометрических базах данных, влекущую за собой низкую цитируемость, могло бы компенсировать создание локальных реферативных баз данных публикаций при научных организациях, по возможности с полными текстами статей. Многие российские журналы продолжают издаваться в печатном виде, не имея электронной версии. Создание и наполнение сотрудниками библиотек электронного списка публикаций своих сотрудников значительно упростило бы доступ к издаваемым статьям. Проведенные исследования доказывают увеличение цитируемости автора в 3 – 5 раз, если к его статьям имеется свободный доступ [2, 3]. Недавние подсчеты выявили практически полное отсутствие репозитариев открытого доступа в России (28) по сравнению с США (301) [16].

Подобная база данных успешно организована и функционирует в ИНГГ СО РАН, где помимо полного списка всех работ сотрудников института за

все годы также была создана собственная программная оболочка [15]. В базе данных присутствуют библиографические описания 40 тысяч публикаций сотрудников, которые находятся в фондах библиотеки института и представлены на его сайте. База создавалась в среде автоматизированной библиотечно-информационной системы CDS/ISIS и включает описания монографий, диссертаций, авторефератов диссертаций, статей в научных журналах, электронных публикаций в Интернет, докладов на конференциях и других публикаций. Более чем для 12 % описаний имеются ссылки на электронные версии публикаций. Важным моментом является интеграция этой базы данных с Web of Science, Scopus и РИНЦ. Для этого был разработан программно-технологический комплекс, позволяющий снабжать метаданные соответствующими ссылками. К описаниям документов, имеющих отражение в вышеуказанных базах данных, добавляются соответствующие идентификаторы записи на эти БД, точно указывающие на соответствующую запись. Такая технология доработки записей позволяет в оперативном режиме отслеживать публикуемость и цитируемость публикаций в перечисленных базах данных. Возможна также интеграция с базой данных Google Scholar. Библиотечный работник имеет возможность выставить для индексирования роботами Google данные своей внутренней базы данных в статических html-страничках, снабдив соответствующие метатэги необходимой информацией о публикациях согласно описанию метаданных Дублинского ядра.

## **2. Редактирование профилей научных организаций**

Специфической в отношении российских учреждений стала проблема графической фиксации российских публикаций в зарубежных базах данных, поскольку они оказались в гораздо большей степени уязвимыми при автоматической обработке в сравнении с изначально англоязычными статьями. В первую очередь это связано с неоднозначностью транслитерации (существует три различных системы перевода кириллицы и латиницу) и сменой названий научных учреждений. Отдельно стоит упомянуть двойную индексацию одной и той же статьи в ее оригинальной и переводной версиях в некоторых базах данных. Эти проблемы чисто технического характера имеют значительные негативные последствия представления информации о российской науке в мировых базах данных Web of Science и Scopus. В частности, публикации ученого или организации оказываются рассредоточенными по нескольким профилям (до нескольких десятков), т. е. идентификаторам в базе данных, аккумулирующим информацию о местах работы авторов, количестве и качестве публикаций, годах публикационной активности, области исследований, соавторах, индексе Хирша, списке использованных в работах литературных источников. С другой стороны, профили могут вообще отсутствовать.

На данный момент еще не сформировалось окончательного мнения насчет того, кто должен заниматься этой работой: сами авторы статей, служба ученого секретаря, технические службы реферативных баз данных, библиотекари или некие специально создаваемые подразделения. На наш взгляд, работники библиотеки наиболее эффективно и быстро способны решить задачу контролирования информации в наукометрических базах данных (заметим, что треть Консультативного комитета Scopus – библиотекари). Это продиктовано спецификой работы с публикациями, владением актуальной информацией об исходящем из научной организации документопотоке, который никто лучше библиотекарей не сверит с тем, как он отражен в базах данных. Учитывая растущий спрос на подобную информацию, новая для библиотекарей роль экспертов в этом деле могла бы привлечь и новых посетителей, количество которых пока неуклонно снижается. Работа библиотекарей в этом направлении может быть важна еще и потому, что, не участвуя напрямую в создании, реферировании и публикации статей, они не имеют материальных и иных выгод, отчего их взгляд на происходящее особенно ценен.

Являясь связующим звеном между создателями, распространителями и потребителями научной информации, библиотекари способны влиять на любое из звеньев, контролировать процесс и занимать значительную позицию при работе с документами. По мнению специалистов ГПНТБ России, именно библиотеки и информационные центры являются независимыми и междисциплинарными учреждениями, способными взять на себя поставку библиометрических данных, необходимых для управления наукой [7], и именно специалисты в библиотеках изучают как рынок научных публикаций, так и потребности ученых [6].

На базе ИНГГ СО РАН и ГНЦ ВБ «Вектор» были разработаны методы обработки информации во внешних реферативных базах данных, а также произведено их аналитическое сравнение [10]. В базе данных Scopus была проведена полномасштабная отработка профилей как обеих организаций в целом, так и профилей каждого из авторов этих организаций в отдельности. В результате этой работы в указанной базе данных любому заинтересованному пользователю в настоящее время доступны наиболее полные и точные данные о публикационной активности наших институтов.

## **3. Оптимизация подписного репертуара**

Повышению числа российских публикаций в зарубежных рейтинговых журналах мог бы способствовать более полный доступ ко всей требуемой информации, для чего необходимо оптимизировать электронную подписку. В литературе была отмечена прямая связь между чтением, доступностью информации и публикационной активностью. Одно из первых исследований было проведено издательством Эльзевир, отметившим, что увеличение числа

доступных статей с появлением электронных версий журналов значительно повысило написание учеными новых работ, а также цитирование. Ту же тенденцию наблюдают сотрудники консорциума НЭИКОН, отмечая прямую взаимосвязь количества скачанных статей с числом опубликованных [14]. Отметим также исследование В. Пислякова и Е. Любушко, проведенное на базе Института катализа СО РАН, в котором рассматривается корреляция между чтением, цитируемостью и созданием оригинальных статей [11].

Проблема заключается в том, что в российских НИИ подписка не всегда соответствует действительным потребностям ученых. В качестве подтверждения укажем на тот факт, что пристатейные списки литературы по естественным специальностям в 2 – 3 раза меньше зарубежных. Используется лишь доступная в конкретной организации литература, которая не всегда коррелирует с международным ядром научных журналов по той или иной тематике.

Особое внимание нужно уделять документам открытого доступа, возникшим как ответ на дороговизну журнальной периодики. Документы открытого доступа представлены бесплатными реферируемыми журналами и архивами электронных препринтов. Сейчас крайне редко можно встретить продуманные системы навигации по этим ресурсам, хотя отдельные исследователи активно ведут работы в этом направлении [8]. Вполне вероятно, что такие архивы могут содержать информацию, способную конкурировать с платной и недоступной в конкретной организации. Изучение этого относительно нового сегмента рынка информации должно быть отдельно учтено при разработке модели комплектования, особенно ввиду ограниченного финансирования научных библиотек. На настоящий момент результаты наших данных [4] показывают, что российские исследователи мало цитируют (а следовательно, мало читают) журналы открытого доступа.

В ИНГГ СО РАН и ГНЦ ВБ «Вектор» были разработаны и прошли апробацию принципиально новые методы формирования ядра научных журналов, основанные на результатах цитат-анализа и анализа ключевых слов из статей сотрудников организации [9, 4, 5].

Информационные потребности пользователей и тенденции их развития, которые можно применять при моделировании репертуара периодики, хорошо выявляются методом цитат-анализа, который мы применили при рассмотрении пристатейной литературы на примере публикаций ученых наших организаций. Тематические профили наших организаций – медико-биологические исследования и геофизика – занимают ключевое место в журнальной периодике, а обслуживающие эти отрасли журналы обладают одними из наиболее высоких импакт-факторов. Поэтому наши методы и результаты вполне могут экстраполироваться на другие научные дисциплины.

Для анализа мы рассмотрели все статьи наших сотрудников за последние пять лет. Из пристатейных списков литературы были выгружены

1300 и 1000 наименований научных журналов в области биологии и геологии соответственно. Доля отечественных изданий составила четверть. В совокупности ученые каждого из институтов сделали порядка 8000 ссылок. По каждой из дисциплин мы получили ранжированные списки журналов, в которых четко прослеживалось ядро.

Другой моделью выявления ядерных журналов является использование ключевых слов из статей сотрудников научной организации в качестве поискового запроса. Не вызывает сомнений тот факт, что публикации научной организации наиболее широко и достоверно отражают реальные достижения организации и сферу научных интересов авторов. Дополнительный список ключевых слов (KeyWords Plus [1]), который присваивается каждой публикации в реферативной базе данных, позволяет сформировать сложный запрос, адресуемый той же базе данных. Список журналов, полученных на основе такого запроса, составленного из машинных ключевых слов, будет напрямую отражать те научные интересы, которые авторы конкретной организации выразили в своих статьях. В нашем случае мы получили 800 журнальных названий в области биологии и 700 – в науках о Земле. Ранжирование журналов позволило четко определить необходимое для подписки ядро.

Результаты обоих методов были успешно применены при корректировании репертуара научной периодики, выписываемого библиотеками двух институтов.

## Заключение

На ближайшие несколько лет перед российской наукой и сферой высшего образования, а также обслуживающими их структурами поставлена масштабная цель упрочения своих позиций на мировой научной арене. Включение научных библиотек в совместное решение задач, связанных с достижением этой цели, могло бы более успешно и эффективно пройти намеченный путь, а также повысить значимость самих библиотек в современной информационной среде. У научных библиотек имеются все возможности и инструменты для формирования внутренних баз данных публикаций, для редактирования профилей организаций во внешних базах данных, а также для поисков новых методов информационного обслуживания российских ученых, что в совокупности могло бы способствовать повышению публикационной активности и цитируемости российских научных и образовательных организаций.

## Литература

- [1] Garfield E. KeyWords Plus – ISI's breakthrough retrieval method. 1. Expanding your searching power on current-contents on diskette // Current Contents. – 1990. – V. 32. – P. 295–299.
- [2] Harnad S., Brody T. Comparing the Impact of Open Access (OA) vs. Non-OA Articles in the Same Journals [Электронный ресурс] // D-Lib

- Magazine. – 2004. – V. 10. – № 6 (June). – Режим доступа:  
<http://www.dlib.org/dlib/june04/harnad/06harnad.html> (Дата обращения: 11.07.2013).
- [3] Lawrence S. Free online availability substantially increases a paper's impact // *Nature*. – 2001. – V. 411. – P. 521.
- [4] Гуреев В. Н., Мазов Н. А. Изучение информационных потребностей ученых с использованием библиометрического анализа для оптимизации комплектования // *Библиосфера*. – 2012. – № 4. – С. 57–66.
- [5] Гуреев В. Н., Мазов Н. А. Моделирование комплектования научной периодики на основе объективной тематики публикаций // *Библиотеки и информационные ресурсы в современном мире науки, культуры, образования и бизнеса: 20-я междунар. конф. "Крым 2013"* (8-16 июня 2013 г., г. Судак): Труды конф. - М.: Изд-во ГПНТБ России, 2013. [Электронный ресурс]. – Режим доступа: <http://gpntb.ru/win/inter-events/crimea2013/predprog.pdf> (Дата обращения: 11.07.2013).
- [6] Земсков А. И., Евстигнеева Г. А. Роль библиотек на мировом рынке научных публикаций // *Вестник РФФИ*, 2005. – № 4(42). – С. 51–56.
- [7] Земсков А. И., Шрайберг Я. Л. Электронная информация и электронные ресурсы: публикации и документы, фонды и библиотеки / под ред. Л. А. Казаченковой. – М.: ФАИР, 2007. – 528 с.
- [8] Канн С. К., Босина Л. В. Развитие навигатора открытых зарубежных научных ресурсов SCIGUIDE // *Библиосфера*. – 2012. – № 5. – С. 41–43.
- [9] Мазов Н. А. Оценка потока научных публикаций академического института на основе библиометрического анализа цитирования // *Информационные технологии в гуманитарных исследованиях*. – 2011. – № 16. – С. 25–30.
- [10] Мазов Н. А., Гуреев В. Н. Новые научные методы для исследования библиотечной отрасли // *Библиосфера*. – 2012. – № 5. – С. 87–90.
- [11] Писляков В. В., Любушко Е. Э. Анализ научно-информационной деятельности (чтение, публикации, цитирование) ученых института катализа им. Г. К. Борескова СО РАН // *Катализ в промышленности*. – 2007. – № 3. – С. 55–63.
- [12] Проект распоряжения Правительства РФ «Комплекс мероприятий, направленных на увеличение к 2015 году доли публикаций российских исследователей в общем количестве публикаций в мировых научных журналах, индексируемых в базе данных «Сеть науки» (Web of science) до 2,44%». – 2013 [Электронный ресурс]. – Режим доступа: <http://минобрнауки.рф/%D0%B4%D0%BE%D0%BA%D1%83%D0%BC%D0%B5%D0%BD%D1%82%D1%8B/3119/%D1%84%D0%B0%D0%B9%D0%BB/1847/13.02.27-%D1%81%D0%B5%D1%82%D1%8C.%D0%BD%D0%B0%D1%83%D0%BA%D0%B8.pdf> (Дата обращения: 11.07.2013).
- [13] Проект создания «Карты Науки России». Установочное заседание экспертных групп проекта: Материалы для обсуждения. – 2013 [Электронный ресурс]. – Режим доступа: [http://www.lin.irk.ru/new/files/sci\\_ross.pdf](http://www.lin.irk.ru/new/files/sci_ross.pdf) (Дата обращения: 11.07.2013).
- [14] Стерлигов И. Кто больше читает, тот больше публикует [Электронный ресурс] // *Наука и технологии РФ*. – 2011 (30 августа). – Режим доступа: [http://www.strf.ru/material.aspx?CatalogId=221&\\_no=42023](http://www.strf.ru/material.aspx?CatalogId=221&_no=42023) (Дата обращения: 11.07.2013).
- [15] Труды сотрудников ИНГГ и ИГМ СО РАН [Электронный ресурс]. – Режим доступа: <http://ibc.ipgg.nsc.ru> (Дата обращения: 11.07.2013).
- [16] Филозова И. А. Открытые архивы научной информации [Электронный ресурс] // *Системный анализ в науке и образовании*. – 2010. – № 1. Режим доступа: [www.sanse.ru/download/49](http://www.sanse.ru/download/49) (Дата обращения: 11.07.2013).

## **Bibliometric methods in development of scientific organization rating**

Vadim N. Gureyev, Nikolay A. Mazov

The article describes the possibilities of scientific libraries to increase scientific organization or university rating in international bibliometric databases, such as Web of Science and Scopus. The authors studied three problems sets that can enable to enhance representativeness of information about Russian scientific organizations and universities in international scientometric databases. At first, we mean designing of home-based abstract databases providing information about publications of authors of institute that makes it possible to perceive specificity of research field and achievements of organization. At second, librarians can directly edit profiles of their scientific organizations in international bibliometric databases, including Web of Science, Scopus, and Google Scholar. Finally, it is possible to develop new methods to provide cutting-edge scientific information to researchers in order to encourage them to create new articles that will enlarge publication indicators in databases. We describe our own experience in using these methods on the base of two scientific institutes: IPGG SB RAS and SCR VB “Vector”.