

УДК 002.1:004.738.5

А. СКАЛАБАН, И. ЮРИК

СИСТЕМЫ АВТОРСКОЙ ИДЕНТИФИКАЦИИ КАК ИНСТРУМЕНТЫ ПОВЫШЕНИЯ ВИДИМОСТИ НАУЧНЫХ ПУБЛИКАЦИЙ В ИНТЕРНЕТ

Научная библиотека, Белорусский национальный технический университет

Рассмотрены некоторые проблемы, связанные с корректностью и полнотой исходных данных при измерении и оценке научной деятельности библиометрическими методами с использованием наукометрических баз данных. Приведен обзор основных систем идентификации авторов. Предложены рекомендации по устранению выявленных проблем и продвижению авторов в мировое научно-информационное пространство.

Введение

В настоящее время в Республике Беларусь для анализа и оценки результатов научной деятельности наряду с традиционными методами – экспертными оценками (отзывами, рецензиями, мнениями экспертов, рецензентов, научных редакторов и др.), все активнее используются библиометрические индикаторы, связанные с учетом числа публикаций и их цитируемостью. Библиометрические данные используются в научной политике и управлении финансированием науки; в национальных программах развития науки и национальных системах оценки результатов научных исследований; при составлении международных и национальных рейтингов вузов и др.

Данные библиометрического анализа позволяют сделать значимые (обоснованные) сравнения научных организаций, дисциплин и стран, но лишь в том случае, если они используются корректно. О необходимости грамотной и аккуратной трактовки наукометрических индикаторов, о проблеме «злоупотреблений цитированием» еще в 60-е годы прошлого века писал один из основоположников библиометрии и наукометрии Ю. Гарфилд. Это проблема не теряет своей актуальности и сегодня. «С одной стороны, это связано с легкостью манипулирования массмедиа этими показателями, с другой стороны – с безответственностью малокомпетентных «специалистов» и администра-

торов науки, с легкостью берущихся за выполнение такой оценки» [1].

Именно в ответ на «игры с цитированием» в 2010 году появился манифест альметрики, предлагающий способы оценки публикаций, отличные от анализа цитирований. Затем в декабре 2012 года на свет появилась декларация по оценке науки – Declaration on Research Assessment (DORA), принятая Американским обществом клеточной биологии (The American Society for Cell Biology (ASCB)). К настоящему времени под ней уже поставили свои подписи 10 668 ученых и 467 организаций [2]. И, наконец, в апреле 2015 года представителями наукометрического сообщества был разработан Лейденский манифест по наукометрии, провозглашающий 10 принципов «в оценке исследовательской деятельности, основанной на наукометрии, с тем, чтобы ученые могли бы проверять тех, кто их оценивает, а «оценщики» могли бы проверять свои индикаторы» [3]. Полагаем, что знакомство с данными документами будет полезно всем представителям отечественного научного сообщества.

В рамках нашей статьи мы хотели бы обратить внимание на проблемы, связанные с корректностью и полнотой исходных данных при измерении и оценке научной деятельности библиометрическими методами, поскольку проблема учета и оценки результатов научной деятельности находится сегодня в фокусе интере-

сов не только государства, в лице чиновников от науки; руководителей научных организаций; самих ученых; но и организаций и исследователей, профессионально специализирующихся в области наукометрии, сбора, систематизации и распространения научной информации, в том числе академических и научных вузовских библиотек.

В настоящее время для библиометрического анализа используются наукометрические базы данных (индексы цитирования) и веб-сервисы, позволяющие в автоматизированном режиме проводить комплексную обработку данных. Наиболее известные среди них: Web of Science (далее – WoS) (производитель Thomson Reuters), Scopus (Elsevier), Google Scholar (Google), Российский индекс научного цитирования (далее – РИНЦ) (Научная электронная библиотека eLIBRARY. RU), а также такие веб-инструменты, как InCites (на основе WoS) и SciVal (на основе Scopus).

Индексы цитирования позволяют осуществлять автоматизированный сбор и анализ библиометрических показателей и, как следствие, дают возможность автоматического расчета соответствующих индексов. «Востребованность их объясняется, с одной стороны, желанием самих ученых вывести «магическую цифирь», сводящую в одном или нескольких числах все таланты и заслуги соратников по цеху, а с другой – стремлением администраторов от науки получить объективные показатели ценности результатов деятельности ученого, научной или образовательной организации или научно-направленного направления. Но не так все просто [4]».

За кажущейся легкостью расчета показателей научной результативности с помощью баз данных научного цитирования скрываются многочисленные подводные камни. За рамками нашей статьи мы оставим те, которые связаны с особенностями публикаций и цитирований в разных научных областях, как, например: практика соавторства и способы счета статей, написанных несколькими авторами; распределение по типам публикаций; различный уровень цитирования; хронология распределения ссылок и др. В данном контексте лишь приведем мнение специалистов в области наукометрии: «...по таким стандартным показателям, как среднее цитирование публикаций или индекс Хирша и его вариации, проводить срав-

нение результативности научной деятельности научных коллективов или отдельных авторов, работающих в разных научных направлениях, представляется весьма затруднительным» [1]. Цель нашей статьи – озвучить некоторые проблемы идентификации данных научных публикаций в индексах цитирования и определить пути их решения.

Проблемы идентификации данных в индексах цитирования и пути их решения

Поскольку различные источники информации о публикациях для оценки научной деятельности формируются по различным принципам, необходимо указывать на основе какого инструмента производится библиометрический анализ. Следует иметь в виду, что показатели, рассчитанные с использованием одной и той же методики, в разных источниках будут иметь разные значения, поскольку расчет в каждом конкретном ресурсе осуществляется на основе информации (источников), содержащихся именно в данном ресурсе (табл. 1).

Так, WoS и Scopus индексирует наиболее значимые в мире журналы в каждой области знаний, труды конференций и, с недавнего времени, книги; РИНЦ – российскую научную периодику и ряд научных журналов стран СНГ; Google Scholar – данные из открытых источников информации (репозитории, журналы открытого доступа и др.). Отличается и глубина архивов (с 1900 г. в WoS, с 1823 Scopus и приблизительно 2000-й – в РИНЦ).

Основное отличие WoS и Scopus заключается в том, что в Scopus при подписке доступ дается ко всему архиву, тогда как ретроспектива баз данных WoS зависит от суммы оплаченных средств, что напрямую влияет на отображаемые показатели.

Международные индексы цитирования WoS и Scopus – это библиографические/реферативные базы данных, в которых собирается и обрабатывается полная библиографическая информация о научных статьях, аннотации/рефераты и пристатейные списки цитируемой в статьях литературы. Полных текстов статей в указанных базах данных нет. В качестве отличительной характеристики РИНЦ и Google Scholar можно привести наличие полных текстов.

В Беларуси лицензии на доступ к ним имеют: WoS – Центральная научная библиотека

Таблица 1. Сравнительный анализ наукометрических инструментов

Инструменты	Массив индексируемых данных	Вид доступа	Глубина архива	Отличительная характеристика	Способы анализа и оценки
Международный индекс цитирования Web of Science (WoS) (Thomson Reuters, 2002)	Более 12000 наиболее значимые журналы в каждой области знания, труды конференций, незначительное количество книг	Подписная база данных, для доступа к которой необходима платная подписка той организации, к которой вы относитесь	с 1900 г.	Ретроспектива БД в конкретной организации зависит от суммы оплаченных средств, что напрямую влияет на отображаемые показатели	Учет числа публикаций, цитирований
Международный индекс цитирования Scopus (Elsevier, 2004)	Более 20000 научных и отраслевых журналов, труды конференций, книжные серии, монографии	Подписная база данных, для доступа необходима платная подписка к той организации, к которой вы относитесь	с 1823 г.	Доступ ко всему архиву, показатели рассчитываются, учитывая журналы с 1996 года, для некоторых журналов крупных издательств – с 1970	Учет числа публикаций, цитирований
Российский индекс научного цитирования (РИНЦ) (Научная электронная библиотека eLIBRARY.RU, 2005)	Научные публикации России и стран СНГ	Доступ бесплатный	с 2000 г.	Наличие полных текстов.	Учет числа публикаций, цитирований
Google Scholar (Google, 2004)	Научно-образовательные ресурсы, доступные в Интернет в репозиториях, журналах открытого доступа, списках процитированной найденной в статьях литературы и др.	Доступ бесплатный	нет данных	Наличие полных текстов	Учет числа публикаций, цитирований

им. Я. Коласа НАН Беларуси; Scopus – Центральная научная библиотека им. Я. Коласа НАН Беларуси, БелСХБ, Научная библиотека БНТУ.

Существенной проблемой, затрудняющей корректный подсчет публикаций и цитирований отдельного автора или научного коллектива в WoS и Scopus является отсутствие однозначной идентификации объектов. В каких случаях могут быть потеряны данные:

- при множественности написания фамилии автора (использование различных схем транслитерации; написание фамилии автора с различным количеством инициалов, смена женских фамилий после замужества, наличии опечаток и пр. [5]);

- при отсутствии в статье данных о месте работы автора либо при наличии нескольких вариантов наименования организации, к которой относится автор (в поле «affiliation» название организации, ее адрес переносятся автоматически).

Как следствие – погрешности в количестве публикаций, цитирований, и соответственно значениях индекса Хирша и др., некорректные итоговые сведения о научных достижениях автора/научного коллектива/региона.

Проблема идентификации данных существует также в РИНЦ и Google Scholar, они несколько отличаются от зарубежных баз данных, на них остановимся далее.

Во многом эти проблемы решаются с помощью уникальных идентификаторов, существующих в виде отдельных инструментов или встроенных в базы данных.

Существует некоторое количество таких систем. Например, Thomson Reuters создан специальный инструмент объединения и корректировки данных автора, связанный с WoS, – авторский профиль **ResearcherID**; в Scopus авторский профиль **Author ID** формируется автоматически для авторов, опубликовавших хотя бы 1 статью; **ORCID** (Open Researcher and Contributor ID) – самостоятельная система – реестр уникальных идентификаторов ученых, независимый от научных дисциплин и национальных границ, синхронизируется с другими системами идентификации, например, с **ResearcherID** и **Author ID**; в РИНЦ создана специальная аналитическая надстройка **SCIENCE INDEX**; в Google Scholar – персональный профиль автора.

Однозначно определить все свои публикации может только сам автор, поэтому настоя-

тельно рекомендуется ученым регистрироваться в системах идентификации и работать со своими персональными профилями. Кроме этого, регистрация авторов в системах идентификации значительно улучшает видимость их работ в мировом научно-информационном пространстве.

Рассмотрим подробнее каждую из приведенных ранее систем идентификации.

ResearcherID – это свободное, общедоступное интерактивное пространство для создания индивидуального номера и персонального профиля. Зарегистрироваться в системе можно на одноименном сайте <http://www.researcherid.com>. Введенные сведения не подвергаются модерации. Не требуется регистрации в системе в том случае, если автор относится к организации, имеющей лицензию на доступ к WoS, и ранее был зарегистрирован в ней. Преимущество регистрации в ResearcherID:

- создание полного списка своих публикаций тремя способами: 1) путем поиска в WoS и сохранения результата; 2) посредством загрузки из библиографического менеджера EndNote; 3) импорт внешних файлов в формате RIS, содержащих статьи, отсутствующие в WoS;
- исключение из своего списка публикаций однофамильцев;
- поиск в WoS по номеру профиля ResearcherID и нахождение всех индексированных публикаций автора, даже при наличии ошибок и разных вариантов написания его фамилии;
- определение наукометрических показателей (анализ цитирований автора будет проводиться только по статьям, проиндексированным в WoS);
- поиск партнеров и др.

Потенциал **ResearcherID** пока не оценен в должной мере белорусскими исследователями, об этом свидетельствуют статистические данные, полученные нами в результате анализа этой системы. Так, по состоянию на 1.12.2014 в **ResearcherID** зарегистрировано 155 пользователей из Беларуси, для сравнения – Россия представлена 32330 пользователями, Украина – 1130, Казахстан – 1869, Польша – 7690, США – 66614. Статистика по научным организациям представлена в табл. 2.

Уникальный идентификатор автора **Scopus Author ID** присваивается при опубликовании даже одной статьи, индексируемой Scopus. Он

Т а б л и ц а 2. Статистика зарегистрированных пользователей в ResearcherID по научным организациям Беларуси

Название организации	Количество пользователей
Belarus State University	29
Belarus State Technological University, BSTU	9
B. I. Stepanov Institute of Physics	8
Belarusian National Technical University	7
Belarusian State University of Informatics and Radioelectronics	6
Brest State Technical University	5
Belarusian Medical Academy of Postgraduate Education	4

выполняет функцию независимого от статьи источника информации, аккумулируя информацию о местах работы автора, количестве его публикаций и их цитируемости, годах публикационной активности, области исследований, соавторах, индексе Хирша, списке использованных в работах литературных источников и пр. Прямое доступа к редактированию своего профиля автор не имеет, вместе с тем уточнить свою библиографическую информацию он может. При наличии доступа к базе данных через свой профиль в Scopus, а при его отсутствии через ссылку <http://www.scopusfeedback.com/>, автор может самостоятельно сделать следующее:

- объединить несколько своих профилей в один;
- уточнить информацию на уровне статей, отмечая свои и исключая чужие публикации в процессе работы со списком;
- выделить все свои статьи и послать запрос по электронной почте в техническую службу Scopus.

Подобную работу может провести не только сам автор. Например, в Белорусском национальном техническом университете (далее – БНТУ) сотрудники Научной библиотеки осуществляли поиск и анализ множественных профилей университета и работающих в нем авторов. Хорошим подспорьем в этой работе стала генерируемая библиотекой база данных трудов научных работников БНТУ. В результате анализа данных в Scopus было обнаружено 27 вариантов названий БНТУ и 170 статей авторов, работающих в университете, но не связанных с его профилем. Был отправлен запрос в службу поддержки. Профили были объединены. В итоге количество статей БНТУ в Scopus увеличилось на 700.

ORCID (Open Researcher and Contributor ID) – совместная разработка нескольких издательств, университетов, научных сообществ. Адрес входа в систему – <http://orcid.org/>.

Основной целью системы является – решить проблему идентификации ученых с одинаковыми именами и фамилиями. **ORCID** обеспечивает следующие функции:

- получение уникального идентификатора и ведение соответствующей учетной записи об исследовательской деятельности;
- программное обеспечение для межсистемной коммуникации.

Следует отметить, что данный номер от авторов последнее время часто требуют издатели во время отправки рукописи в журнал. Таким образом, этот номер, отсылающий к отредактированному списку публикаций автора, является своеобразной визитной карточкой автора, поскольку информация видна редактору и (иногда) рецензентам.

Учетная запись **ORCID** включает в себя информацию об имени ученого, его электронном адресе, названии организации и его исследовательской деятельности. **ORCID** учитывает необходимость контроля за распространением этих данных и предоставляет соответствующие инструменты для управления уровнем приватности данных.

ID представляет собой 16-значный числовой код, согласованный со стандартом ISO (ISO 27729). **ORCID ID** – это URI, поэтому отображается как адрес вида <http://orcid.org/xxxx-xxxx-xxxx-xxxx>.

Записи в **ORCID** хорошо синхронизируются с записями в базе данных Scopus: во-первых, можно осуществлять перенос всех публикации автора из Scopus в **ORCID**, указав номер авторского профиля **Scopus Author ID**; во-вторых, если работы автора в Scopus присоединены к нескольким профилям, то при включении их в **ORCID** служба технической поддержки Scopus отредактирует профиль автора в базе данных.

Согласно принципам организации **ORCID**, участником этого проекта может стать любая организация научного сообщества. Для индивидуальных исследователей доступ к реестру **ORCID** предоставляется бесплатно. По состоянию на 1.12.2014 г. в системе было зарегистрировано 1 млн. пользователей и содержалось 5359470 работ.

SCIENCE INDEX – специальная аналитическая надстройка в РИНЦ, с помощью которой производится редактирование автором своих личных данных, а также анализ публикационной активности и цитируемости. Приводим данные статистики. По состоянию на 1 декабря 2014 г. всего в РИНЦ зарегистрировано 276491 авторов, из них 266632 – из России; 3671 – из Украины; 3397 – из Беларуси; 1096 – из Казахстана. При этом поиск по автору и стране (Беларусь) в РИНЦ дает результат – 9037.

В отличие от WoS и Scopus в РИНЦ нет проблемы с авторским множеством, однако есть другая – часть статей может быть не привязана к профилю автора. Анализ библиометрических показателей 10 самых цитируемых в РИНЦ белорусских авторов показал, что из 10 только 6, отмеченных символом звездочка «*» зарегистрированы в **SCIENCE INDEX**. Вполне вероятно, что сведения о публикациях и цитировании авторов, не прошедших регистрацию в РИНЦ, не полные.

Пищу для размышления сотрудникам и администрации учреждений высшего образования Беларуси (УВО) дает и анализ количества зарегистрированных в **SCIENCE INDEX** авторов по вузам (табл. 3).

Таблица 3. Топ-10 УВО Беларуси по числу зарегистрированных в **SCIENCE INDEX** авторов

N п/п	Название организации	Количество зарегистрированных авторов
1	Белорусский государственный медицинский университет	635
2	Гродненский государственный медицинский университет	443
3	Витебский государственный университет им. П. М. Машерова	294
4	Белорусская медицинская академия последипломного образования	207
5	Витебский государственный медицинский университет	199
6	Белорусский государственный университет	196
7	Витебская государственная академия ветеринарной медицины	131
8	Белорусский государственный технологический университет	64
9	Гродненский государственный университет им. Янки Купалы	57
10	Белорусский национальный технический университет	48

Улучшению представления научных результатов белорусских научных и образовательных организаций в РИНЦ могли бы способствовать следующие тактические шаги: 1) обязательная регистрация каждого публикуемого сотрудника в SCIENCE INDEX; 2) активная работа авторов, прошедших регистрацию, со своими списками – выявление «непривязанных» статей, поиск своих статей в приставной литературе и проч.; 3) расширение ассортимента белорусских научных журналов на платформе elibrary.ru (как по названию, так и по глубине архивов).

С точки зрения реальных показателей цитируемости для авторов из белорусских научных и образовательных учреждений Google Scholar представляет больший интерес, чем WoS или Scopus, поскольку система индексирует массив публикаций, размещенных в институциональных репозиториях (на 1.06.2015 в OpenDOAR зарегистрировано 15 репозиторий УВО Беларуси, объем документов – более 146000), а также другие источники на русском языке (научные журналы открытого доступа, материалы конференций, препринты и др.). Вместе с тем следует отметить, что эта система до сих пор находится в стадии бэта-версии, поэтому практически не может использоваться для официальных отчетов (лишь иногда ее включают как факультативную).

Для получения данных о публикациях для оценки научной деятельности с помощью Google Scholar авторам необходимо создавать личный профиль в этой системе, который, ко всему прочему, позволит сохранять результаты поиска; следить за цитированием интересующих работ; отслеживать отдельные работы или творчество отдельных авторов.

Вход в систему по адресу: <http://scholar.google.ru/>. Сначала необходимо создать новый аккаунт в Google или войти в уже существующий. Заполнение формы регистрации в Google Scholar осуществляется в три шага:

- 1 шаг – заполнение полей – имя, место работы, электронная почта для подтверждения, области интересов;
- 2 шаг – работа со статьями – необходимо указать именно свои статьи в списке, автоматически предложенном системой, а также можно добавлять публикации самостоятельно;

- 3 шаг – заполнение страницы профиля – установка параметров обновления профиля, загрузка фотографии и др.

Особо следует отметить функцию экспорта списка своих работ в другие системы идентификации авторов, как, например, **ResearchID** и **ORCID**. В связи с этим авторам из белорусских научных и образовательных учреждений, чьи публикации размещены в институциональных репозиториях, мы рекомендуем сначала создавать персональные профили в Google Scholar, добавляя в них список собственных статей, а затем их экспортировать в другие системы идентификации авторов.

Преимущество Google Scholar как наукометрического инструмента – максимальная широта представленных материалов. Недостатки – в нее не попадают сведения из закрытых источников и публикации, отсутствующие в электронном виде (или упоминания о таковых); нет гарантии в достоверности данных, добавленных в свой личный профиль самим автором. Все зависит от того, насколько он соблюдает нормы научной этики.

Выводы

1. Обязательным условием при измерении и оценке научной деятельности библиометрическими методами является корректность и полнота исходных данных. Необходимо учитывать существующие в наукометрических инструментах проблемы с идентификацией авторов и организаций. Игнорирование может привести к занижению реальных показателей публикационной активности и цитируемости как отдельных авторов, так и научных коллективов в целом.

2. Во избежание ошибок и неточностей при индексировании публикаций научных и образовательных организаций возможно поможет строгая регламентация указания названия организации как места работы на административном уровне.

3. Регистрация в системах идентификации, а затем регулярная работа автора со своим профилем позволяет не только контролировать корректность представленных данных в различных наукометрических инструментах, но и улучшает видимость научных работ.

4. Особенностью современной научной коммуникации является формирование электронного научного пространства, а ее эффектив-

ность становится в большей мере зависимой от активного освоения и использования учеными разных форм электронной научной коммуникации. В связи с появлением альтернативных цитированию способов оценки научных результатов с помощью альметрик научным организациям и их сотрудникам следует принимать во внимание важность качественного присутствия в Интернет: в блогосфере, социальных сетях, в системах идентификации авторов, библиографических менеджерах.

Авторы выражают глубокую признательность М. А. Акоеву (заместителю директора Центра мониторинга науки и образования Уральского федерального университета имени первого Президента России Б. Н. Ельцина) и Н. А. Мазову (к. т. н., заведующему информационно-библиотечным центром Института нефтегазовой геологии и геофизики им. Академика А. А. Трофимука СО РАН, Новосибирск), чьи замечания и предложения были полезны для подготовки настоящей статьи.

Литература

1. **Руководство** по наукометрии: индикаторы развития науки и технологии: [монография] / М. А. Акоев, В. А. Маркусова, О. В. Москалева, В. В. Писляков; [под ред. А. М. Акоева]. – Екатеринбург: Изд-во Урал. ун-та, 2014. – 250 с.
2. **Обновление** инструкции для авторов научных журналов: методические материалы / пер. с англ. А. В. Бажанова; пер. с англ. под ред. А. Ю. Гаспаряна, О. В. Кирилловой. – СПб.: Сев.-Зап. ин-т упр. – фил. РАНХиГС, 2015. – 48 с.
3. **Лейденский манифест** для наукометрии / Институт всеобщей истории Российской академии наук; перевод А. А. Исэрова. – Режим доступа: <http://www.igh.ru/about/news/1053/>
4. **Новиков**, Д. А. Померяемся «Хиршами»? (Размышления о наукометрии) / Д. А. Новиков // Высшее образование в России. – 2015. – № 2. – С. 5–13.
5. **Мазов**, Н. А. Роль единых идентификаторов в информационно-библиографических системах / Мазов Н. А., Гуреев В. Н. // Научно-техническая информация. Серия 1: Организация и методика информационной работы. – 2014. – № 9. – С. 32–37.

Поступила 04.11.15

A. SKALABAN, I. YURIK

AUTHOR'S IDENTIFICATION SYSTEMS AS AN INSTRUMENT FOR ENHANCING OF RESEARCH PUBLICATIONS VISIBILITY ON THE WEB

Belarusian National Technical University, Library

The purpose of the article is to give the reader some information on some problems related to the correctness and completeness of the initial data in the measurement and evaluation of scientific activity by bibliometric methods using scientometric databases. The article gives a detailed analysis of the main systems of identification of the authors. Recommendations to address the identified problems and the promotion of authors in the world scientific information space are described in the article.