

ЕЛЕНА ВАСИЛЬЕВНА ФОТИНАведущий программист Регионального центра новых информационных технологий, Петрозаводский государственный университет (Петрозаводск, Российская Федерация)
*fotina@petsu.ru***АЛЕКСЕЙ ГЕОРГИЕВИЧ МАРАХТАНОВ**заведующий сектором научно-образовательных Интернет-систем Регионального центра новых информационных технологий, Петрозаводский государственный университет (Петрозаводск, Российская Федерация)
*marahthanov@petsu.ru***ОЛЬГА ЮРЬЕВНА НАСАДКИНА**кандидат технических наук, заместитель директора Регионального центра новых информационных технологий, Петрозаводский государственный университет (Петрозаводск, Российская Федерация)
onasad@petsu.ru

ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ПОВЫШЕНИЯ ПОКАЗАТЕЛЕЙ ИНДЕКСОВ НАУЧНОГО ЦИТИРОВАНИЯ УЧЕНЫХ ПетрГУ*

Рассматриваются существующие индексы цитирования, основные методы, применяемые в данной области с целью разработки информационно-методической системы для улучшения показателей индексов научного цитирования ученых ПетрГУ.

Ключевые слова: цитирование, индекс цитирования, индекс Хирша, импакт-фактор

Региональный центр новых информационных технологий Петрозаводского государственного университета с начала 2012 года работает над проектом «Научные основы информационно-аналитического обеспечения реализации единой программной и информационной среды управления научно-исследовательской деятельностью вуза». Одной из задач данного проекта является разработка механизмов мониторинга и улучшения показателей индексов научного цитирования сотрудников ПетрГУ. Для решения обозначенной задачи необходимо исследовать предметную область (принципы формирования индексов научного цитирования), определить место ученых ПетрГУ в этих индексах и адекватность имеющихся показателей, а также возможные причины некорректного расчета показателей, по результатам исследования наметить основные возможности и подходы к улучшению существующих показателей и реализовать их на практике.

Исследование предметной области целесообразно начинать с определения цитируемости. Можно говорить о цитируемости ученого, цитируемости статьи и цитируемости журнала.

Статья А цитирует статью Б, если хотя бы один раз в тексте А имеется ссылка на Б и Б, таким образом, вынесена в А в пристатейный список литературы или фигурирует в постраничной сноске. Журнал J цитирует журнал I столько раз, сколько статей из J цитируют статьи из I [3]. По утверждению Е. Гарфилда [9], частота цитирований отражает влияние определенной статьи или определенного ученого и является мерой

научно-исследовательской активности или мерой коммуникации на основе этой активности. Эта мера является социометрической. Число цитирований некоторой работы ученого само по себе не является мерой значимости и должно использоваться лишь в совокупности с другими оценками, хотя подсчет цитирований более правомерен, чем подсчет публикаций.

Если говорить о цитируемости ученого, то одним из самых известных показателей является индекс Хирша (h-индекс). Ученый имеет индекс h, если h из его Np статей цитируются как минимум h раз каждая, в то время как оставшиеся (Np – h) статей цитируются не более чем h раз каждая [11].

В свою очередь деятельность научных журналов характеризует импакт-фактор. Его расчет основан на данных трехлетнего периода. Например, импакт-фактор журнала в 2011 году I_{2011} вычисляется следующим образом: $I_{2011} = A/B$, где A – число цитирований статей, опубликованных в данном журнале в 2009–2010 годах, в других журналах, отслеживаемых Институтом научной информации (Institute for Scientific Information, ISI), в течение 2011 года; B – общее число статей, опубликованных в данном журнале в 2009–2010 годах. Импакт-фактор полезен для выявления значимости абсолютного (общего) числа цитирований. Он устраняет некоторую предвзятость в подсчетах, когда крупные журналы оказываются в преимуществе перед небольшими журналами, часто издаваемые – перед редко издаваемыми и старые журналы – перед новыми. В этих случаях у таких журналов боль-

шее число цитируемых печатных материалов, чем у журналов менее крупных или более новых. При прочих равных условиях, чем больше ранее опубликованных статей, тем чаще журнал будет цитироваться [15]. Есть разные мнения по поводу использования импакт-фактора и эффективности оценки журналов таким образом. Наиболее лаконичную оценку этому показателю дал Хозффель. Согласно его утверждению, «импакт-фактор не является совершенным инструментом для измерения качества статей, но нет и ничего лучшего, и его преимущество в том, что он уже существует и, следовательно, представляет собой хороший метод научной оценки. Опыт показал, что в каждой специальности лучшими журналами являются те, в которых труднее всего опубликоваться, и это те журналы, у которых высокий импакт-фактор. И многие из этих журналов появились задолго до того, как импакт-фактор был изобретен» [10]. Импакт-фактор применяется уже в течение долгого времени и является самым известным и широко используемым показателем для журналов. В то же время несколько лет назад появился SCImago journal rank (SJR) indicator, который часто сравнивают с импакт-фактором [6]. В качестве источника берутся данные индекса цитирования Scopus. Оба показателя подсчитываются делением числа цитирований в некотором журнале на общее число статей в этом журнале за определенный период. Но SJR индикатор зависит от «престижа» журнала и не учитывает самоцитирования. Для оценки качества научных журналов SJR индикатор применяет PageRank алгоритм, используемый поисковой системой Google для оценки веб-страниц. Также необходимо отметить и тот факт, что, в отличие от импакт-факторов, доступных по подписке, данные SJR индикатора находятся в открытом доступе.

В настоящее время использование показателей цитирования становится все более востребованным, поскольку они являются показателями результативности деятельности вуза, а также представляют индикаторы качества для исследователей, научных организаций и научных периодических изданий. Необходимо отметить, что показатели цитирования зависят от разных факторов, в том числе и от области знаний. В частности, показатели биологов и медиков выше, чем у физиков, а у физиков выше, чем у математиков. Это связано в первую очередь с общим числом ученых в области медицины, физики или математики в целом. С другой стороны, сравнивая таким образом исследователей, следует учитывать и тот факт, что при подсчете показателей учитываются по большей части только журнальные статьи и некоторые публикации конференций [1]. Что касается непосредственно индекса Хирша, то необходимо добавить, что, конечно же, не все статьи ученого способствуют увеличению дан-

ного индекса. Ряд статей с низким числом цитирований никогда не будут учтены при подсчете индекса Хирша. Большинство статей получают наибольшее число цитирований в определенный период времени и затем больше не цитируются. Иными словами, те статьи, которые увеличивали h-индекс ученого в ранний период его карьеры, в более поздний период карьеры учитываться в h-индексе не будут. Но несомненно, индекс Хирша уменьшаться со временем не может [9]. Эти и другие особенности, свидетельствующие о некоей доле относительности различных показателей, следует учитывать, формируя выводы, основанные на их значениях. Показатели цитирования дают возможность в некоторой степени сопоставить уровень научных исследований разных организаций или отдельных исследователей. Поэтому их значения, в частности, оказывают влияние на позиции вуза в различных рейтингах. Например, в Национальном рейтинге университетов (<http://unirating.ru/>) учитываются в том числе и наукометрические данные, такие как общее число публикаций, суммарное число цитирований публикаций организации, индекс Хирша в базах данных цитирования Web of Science, Scopus и РИНЦ.

Значения показателей часто учитываются при определении победителей грантов, конкурсов, программ. Например, при подаче заявки в Российский гуманитарный научный фонд, в частности, нужно указать количество публикаций в изданиях, входящих в системы индексов цитирования «Social Sciences Citation Index» (SSCI), «Arts and Humanities Citation Index» (AHCI) и Российский индекс цитирования (РИНЦ) (http://www.rfh.ru/downloads/IS_RFH.pdf).

Другим важным объектом рассматриваемой области является индекс цитирования. Индекс цитирования – это библиографическая база данных, указатель ссылок между публикациями, позволяющая определить статьи, цитируемые данной публикацией и цитирующие ее. Первый индекс цитирования для статей, опубликованных в научных журналах, был предложен Е. Гарфилдом в 1955 году в статье «Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas» [7]. Гарфилд в этой статье говорит о новом подходе в научной литературе, когда индекс цитирования сводит вместе тот материал, который никогда бы не был сопоставлен обычным предметным индексированием. Он описывает индекс цитирования как индекс соединения идей. Благодаря индексу цитирования автор может без труда определить, какие из ученых ссылались на его работу, что, в свою очередь, способствует большей коммуникации между учеными. Возможно, ученый даже узнает о влиянии, которое оказала его работа и о котором он даже не догадывался. Таким образом, полезность индекса цитирования можно рассматривать с точки зрения передачи идей.

Цитирование предполагает взаимосвязь между частью или всеми цитируемыми статьями с частью или всеми цитирующими статьями. Проследив за последующими цитированиями, можно увидеть историю развития некоторой идеи: где и как она применялась, была ли подтверждена, отклонена или вовлечена в дальнейшую работу. Связь идей обеспечивается ссылками, которые делает автор, на более ранние статьи. Но таким образом устанавливаются лишь ссылки в прошлое. Благодаря индексам цитирования ссылки реорганизуются, обеспечивая дополнительную возможность проследить развитие этих же тем, но уже по времени вперед [12].

Индекс цитирования выполняет две фундаментальные функции [8]. Во-первых, он дает информацию о том, что было опубликовано. Во-вторых, он показывает, как каждый кирпич в здании науки связан со всеми остальными.

В базах данных цитирования хранится полная библиографическая информация о статье, аннотация (если есть) и список цитируемой литературы, а также информация об авторах и местах их работы. Самыми известными международно признанными индексами цитирования являются Web of Science и Scopus.

Web of Science предоставляется компанией Thomson Reuters. На данный момент предлагается поиск среди свыше 12 000 журналов и 148 000 материалов конференций, 50 миллионов записей и 800 миллионов цитирований, охватывая период цитирования начиная с 1900 года. Web of Science включает в себя следующие базы данных: Science Citation Index Expanded, Social Sciences Citation Index, Arts & Humanities Citation Index, Index Chemicus, Current Chemical Reactions и Book Citation Index. Все заголовки журналов, включенных в научные продукты Thompson Reuters, доступны через Master journal list (<http://ip-science.thomsonreuters.com/mjl>). Можно просматривать полный список журналов, а также осуществлять поиск по базе данных, ключевым словам, заголовку журнала или его ISSN. С помощью этого списка можно выбрать значимый журнал для публикации научной статьи или журнал, который представляет наибольший интерес научной общественности в конкретной области знаний. В Thomson Reuters добавление и исключение журналов осуществляется постоянно в течение года. Ежегодно рассматриваются свыше 2000 журналов, и 10–12 % из них включаются в базу. Более того, текущий список журналов, представленных в продуктах Thomson Reuters, постоянно пересматривается. При оценке журналов оцениваются разные факторы: базовые издательские стандарты журналов, их содержание, международный состав авторов и данные по цитированию [2].

Реферативная база данных SciVerse Scopus индексирует более 18 000 наименований на-

учно-технических и медицинских журналов примерно 5 000 международных издательств. Предоставляется компанией Elsevier. База данных Scopus обновляется ежедневно. В Scopus включены журналы из всех географических регионов, в том числе и не англоязычные (но могут идти с аннотацией на английском языке). В общей сложности примерно 21 % журналов в Scopus опубликованы на других языках, кроме английского (или опубликованы и на английском, и на каком-то другом языке). Более чем половина содержимого представлена другими странами Европы, Латинской Америки и Азиатско-Тихоокеанского региона [14].

Необходимо отметить, что вышеуказанные индексы цитирования, Web of Science и Scopus, доступны только по подписке.

Существуют индексы цитирования с бесплатным доступом. Автоматический индекс цитирования Google Scholar (Google Академия) обеспечивает возможность поиска научной литературы в глобальной сети. Таким образом, он охватывает все науки и дисциплины, языки и страны, а также все типы источников публикаций. Google Scholar позволяет выполнять поиск среди прошедших рецензирование научных журналов, статей, тезисов, диссертаций, книг, препринтов, презентаций и технических отчетов, предоставляемых университетами, научными институтами, профессиональными сообществами, исследовательскими группами. Подчас Google Scholar позволяет найти цитирования в журналах и материалах конференций, не включенных в другие индексы цитирования. Одним из важнейших свойств Google Scholar является возможность ученому проследить цитирования вперед по времени: с каждой статьей предоставляется список тех более поздних статей, которые цитируют данную. При поиске наиболее цитируемые статьи будут расположены раньше. Из неудобств использования Google Scholar можно отметить некоторые неточности с подсчетом цитирований, проблемы с идентификацией однофамильцев: ученые, имеющие одинаковые инициалы, не будут различаться в системе, и в данной базе цитирования отсутствует классификация по предметным областям [13].

В 2005 году появился Российский индекс научного цитирования (РИНЦ). Он включает в себя более 2 миллионов публикаций российских авторов, а также информацию о цитировании этих публикаций из более 3000 российских журналов. Основная задача, которая стоит перед РИНЦ, это создание максимально полной базы данных научных публикаций российских ученых. Благодаря отечественному индексу цитирования намного облегчилась задача поиска публикаций российских авторов и поиска информации о самих ученых. Помимо собственно публикаций с их библиографической и цитатной инфор-

мацией, РИНЦ предоставляет аналитическую информацию об авторах. По каждому автору в базе данных РИНЦ приводится анализ публикационной активности, включающий в себя различные показатели по числу цитирований и числу публикаций, а также статистические отчеты по распределению публикаций и цитирований. Доступ и регистрация в РИНЦ бесплатны. Авторы, зарегистрировавшиеся в системе Science Index, являющейся аналитической надстройкой над РИНЦ, получают доступ к целому ряду различных инструментов и сервисов. В частности, для зарегистрированных пользователей на странице с библиографическим описанием публикации показывается, помимо числа цитирований в РИНЦ, и число цитирований данной публикации в Web of Science и Scopus. Также, что особенно важно, зарегистрированные пользователи могут работать со списком своих публикаций, добавляя в него ссылки, которые система не смогла приписать ему автоматически, или, наоборот, удалить те ссылки, которые были приписаны по ошибке. Причины некоторых неточностей в списках публикаций могут быть следующие. Во-первых, в базе данных системы встречаются однофамильцы, причем одинаковыми бывают и инициалы авторов. Во-вторых, довольно часто встречаются ошибки в фамилиях или инициалах авторов. Также подчас автор вообще не включен в список авторов публикации, так как некоторые журналы ограничивают число авторов в списке цитируемой литературы [4], [5].

Ученые Петрозаводского государственного университета представлены в каждом из вышеперечисленных индексов цитирования.

Необходимо отметить некоторые проблемы с неоднозначностью представления данных, особенно в зарубежных индексах цитирования. В частности, название Петрозаводского университета в Web of Science записано в четырех вариантах: «petrozavodsk state univ», «univ petrozavodsk», «petrozavodsk univ» и «petrozavosk state univ». Разные способы написания встречаются и у авторов, что вызвано рядом причин, самой распространенной из которых является сложность транслитерации, например «petrovskiy» и «petrovsky». Помимо этого могут быть указаны полные или неполные инициалы: «petrov v» и «petrov ve». Кроме того, ошибки могут возникнуть после распознавания текста: «pergaunent al» и «pergaunent al» для автора «pergaunent al».

Аналогичным образом в Google Академии результаты поиска публикаций ПетрГУ с указанием названия университета на русском или английском языках будут отличаться. Также и для подсчета числа цитирований, индекса Хирша и числа публикаций Google Академия показывает разные цифры в зависимости от формы записи имени автора: на русском языке или транслитерацией.

Из вышесказанного следует, что для успешного продвижения авторов научных публикаций, повышения их личных показателей цитирования необходимо соблюдение целого ряда условий. Ключевым фактором является написание ученым интересных статей, которые будут публиковаться и цитироваться. Но помимо этого ему необходимо предпринять ряд шагов, которые будут этому способствовать. Как минимум ученый должен знать об имеющихся показателях, на основе которых проводится оценка его научной деятельности, о существующих в настоящее время базах данных цитирования и тех возможностях, которые ему предоставляются после регистрации, персонализации в этих системах. Кроме этого, опираясь на предоставляемые индексами цитирования списки высокоимпактных журналов, ученый будет знать, в каких журналах ему лучше печататься.

Результаты исследования показывают, что частой причиной некорректных данных в индексах научного цитирования о цитируемости статей сотрудников ПетрГУ является несоблюдение авторами формальных принципов подготовки статей к публикации, их оформления, а также выбора соответствующих изданий для размещения статей. Для решения обозначенной проблемы необходима реализация информационно-методической поддержки ученых ПетрГУ – авторов статей, которая должна заключаться в предоставлении своевременной и актуальной информации о существующих индексах научного цитирования, принципах их формирования, об индексируемых научных изданиях, а также о текущем месте в индексах и фактических значениях показателей самого автора статьи.

Одним из способов осуществления информационно-методической поддержки может стать предоставление каждому ученому ПетрГУ доступа к информационной системе, содержащей обозначенные выше сведения. В связи с этим было принято решение о создании информационно-аналитической системы «Показатели цитирования ученых ПетрГУ». Информационная система должна включать в себя следующие основные разделы:

- «Об индексах цитирования»;
- «Инструкции, рекомендации»;
- «Вспомогательные сервисы»;
- «Онлайн-консультант»;
- «ПетрГУ в индексах цитирования» и пр.

В разделе «Об индексах цитирования» будет доступна базовая информация о самых известных индексах цитирования, таких как Web of Science, РИНЦ, Google Академия и Scopus. В модуле «Инструкции, рекомендации» пользователям системы будут, в первую очередь, предложены инструкции по регистрации, созданию своего профиля в основных индексах цитирования. Как уже отмечалось, это дает целый ряд

возможностей для авторов, наиважнейшая среди которых – работа со своим списком публикаций. Кроме этого, создание профиля помогает расширить круг ученых, с которыми автор сотрудничает. Пользователь сможет изучить подробные текстовые инструкции или же посмотреть видеоинструкции, в которых по шагам рассказывается процедура регистрации. Также в этом модуле будут рекомендации по выбору журналов и подготовке научных статей. Наряду с текстовыми рекомендациями будут представлены списки высокоимпактных журналов по оценке таких индексов цитирования, как Web of Science и РИНЦ. В разделе «Онлайн-консультант» пользователям будет доступна форма обратной связи, позволяющая задать любые вопросы по теме индексов цитирования экспертам, например специалистам отдела научных исследований

вуза. Среди сервисов раздела «Вспомогательные сервисы» необходимо выделить предлагаемую пользователям стандартную транслитерацию. Как уже отмечалось, одной из проблем в международных индексах цитирования является неоднозначность записи имен ученых, как результат – могут возникать неточности в привязке публикаций к автору и поэтому производиться неправильные подсчеты показателей. В дальнейшем планируется добавление модуля «ПетрГУ в индексах цитирования», отображающего данные по показателям ученых университета в разных индексах цитирования, а также включающего в себя диаграмму роста показателей ПетрГУ.

Запуск информационной системы в эксплуатацию запланирован на 2013 год. В настоящее время разработан прототип системы, который тестируется группой экспертов.

* Работа выполнена при поддержке Программы стратегического развития (ПСР) ПетрГУ в рамках реализации комплекса мероприятий по развитию научно-исследовательской деятельности на 2012–2016 гг.

СПИСОК ЛИТЕРАТУРЫ

1. Арнольд Д., Фаулер К., Шаракшанэ С. Индекс цитирования – инструмент, а не цель! [Электронный ресурс]. Режим доступа: http://www.nanometer.ru/2011/04/27/nauchnie_publicacii_258790.html
2. Джеймс Теста. Процесс отбора журналов в Thomson Reuters [Электронный ресурс]. Режим доступа: http://thomsonreuters.com/content/science/pdf/ssr/journal_selection_essay-russian.pdf
3. Писляков В. В. Наукометрические методы и практики, рекомендуемые к применению в работе с российским индексом научного цитирования // Отчет о научно-исследовательской работе (промежуточный) по теме «Разработка системы статистического анализа российской науки на основе данных российского индекса цитирования». М., 2005.
4. Российский индекс научного цитирования [Электронный ресурс]. Режим доступа: http://elibrary.ru/project_risc.asp
5. Российский индекс научного цитирования: успехи и проблемы. Проект глазами его разработчиков [Электронный ресурс]. Режим доступа: <http://www.polit.ru/article/2009/01/26/rints/>
6. Falagas M. E., Kouranos V. D., Arencibia-Jorge R., Karageorgopoulos D. E. Comparison of SCImago journal rank indicator with journal impact factor // The FASEB Journal. 2008. Vol. 22. P. 2623–2628.
7. Garfield E. Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas // Science. 1955. Vol. 122, № 3159. P. 108–111.
8. Garfield E. Citation Indexing for Studying Science // Essays of an Information Scientist. 1970. Vol. 1. P. 132.
9. Garfield E. Citation Frequency as a Measure of Research Activity and Performance // Essays of an Information Scientist. 1973. Vol. 1. P. 406–408.
10. Garfield E. The Evolution of the Science Citation Index Search Engine to the Web of Science, Scientometric Evaluation and Historiography, 2007 [Electronic resource]. Access mode: <http://garfield.library.upenn.edu/papers/barcelona2007.pdf>
11. Hirsch J. E. An index to quantify an individual's scientific research output // Proc. Nat. Acad. Sci. 2005. Vol. 102. № 46. P. 16569–16572.
12. Malin M. V. The Science Citation Index, a new concept in indexing // Library Trends. 1968. Vol. 16. № 3. P. 374–387.
13. Noruzi A. Google Scholar: The New Generation of Citation Indexes // Libri. 2005. Vol. 55. P. 170–180.
14. SciVerse Scopus Content Coverage Guide [Electronic resource]. Access mode: http://www.info.sciverse.com/UserFiles/sciverse_scopus_content_coverage_0.pdf
15. The Thomson Reuters Impact Factor [Electronic resource]. Access mode: http://thomsonreuters.com/products_services/science/free/essays/impact_factor/