

УДК 37.013

DOI: 10.34670/AR.2019.44.1.032

Роль наукометрических и библиографических баз данных в сферах науки и образования

Калистратов Дмитрий Сергеевич

Кандидат технических наук, докторант, ассистент,
Тульский государственный университет,
300012, Российская Федерация, Тула, пр. Ленина, 92;
e-mail: Kalistratow@list.ru

Аннотация

Проанализирована роль наукометрических и библиографических баз данных в сферах науки и высшего профессионального образования, раскрыты понятия метрики, метрического пространства и наукометрической базы данных, представлены краткие сведения о наиболее распространенных базах данных Web of Science, Scopus и Erih Plus, указаны их сходства и различия, приведены ссылки на интернет-ресурсы, содержащие справочные сведения о включении периодических изданий в указанные базы данных, с точки зрения объективности и пользы для науки и образования исследованы основные критерии сравнительной оценки периодических изданий, отдельных ученых, а также научных и образовательных учреждений, в частности, раскрыты понятия, касающиеся рейтинга научного издания, изучено понятие импакт-фактора и ранга периодического издания, на конкретных примерах продемонстрированы особенности роста индекса цитирования и индекса Хирша отдельно взятого ученого, дана сравнительная оценка этих критериев, указаны их преимущества и недостатки, по результатам анализа выявлены позитивные и негативные аспекты использования наукометрических и библиографических баз данных в сферах науки и образования, а также указана роль наукометрических баз данных для науки, образования и общества в целом.

Для цитирования в научных исследованиях

Калистратов Д.С. Роль наукометрических и библиографических баз данных в сферах науки и образования // Педагогический журнал. 2019. Т. 9. № 1А. С. 87-94. DOI: 10.34670/AR.2019.44.1.032

Ключевые слова

Наука, образование, педагогика, наукометрическая база данных, преимущества, недостатки, импакт-фактор, ранг издания, индекс цитирования, индекс Хирша.

Введение

В последнее время особую актуальность для ученых и педагогов приобрели научные базы данных Web of Science, Scopus, Erih Plus [Web of Science, www; Scopus, www; ERIH PLUS, www]. Например, в Российской Федерации необходимым условием для допуска к защите диссертации на соискание ученой степени доктора наук по технической специальности обязательным требованием сейчас является наличие как минимум двух публикаций, индексируемых в базе данных Scopus. Не меньшую роль наличие таких публикаций играет и для преподавателей высших учебных заведений, поскольку свидетельствует о востребованности и квалификации специалиста в сфере образования.

Таким образом, престиж и научный «вес» публикации указанного плана на «внутренней» и международной арене очевиден. Однако, молодым ученым и специалистам в области науки и образования, впервые столкнувшимся с понятием наукометрической базы данных, бывает сложно сразу разобраться во всех тонкостях и нюансах, сопутствующих терминологии и принципу работы этих баз, а также сформировать более или менее объективное впечатление обо всех преимуществах и недостатках организации их деятельности.

Метрика и метрическое пространство

По указанным причинам, в первую очередь поясним, что такое метрика и что такое метрическое пространство. Если подходить к определению метрики строго формально, то это слово означает правило, которое позволяет определять разницу (расстояние, рассогласование, различие) между двумя любыми элементами соответствующего метрического (измеримого) пространства. Таким образом, наукометрическая база данных – это база, которая позволяет проводить численную оценку и сравнение показателей научной эффективности элементов рассматриваемого метрического пространства.

Например, в метрическом пространстве научных журналов метрикой может являться разница импакт-факторов двух журналов, говорящая в пользу того журнала, импакт-фактор которого больше. В метрическом пространстве ученых метрикой может служить разница индексов цитирования, либо индексов Хирша. Однако, сразу же возникает вопрос о том, всегда ли объективна данная численная оценка. Прежде чем ответить на этот вопрос, кратко опишем наиболее распространенные базы данных.

Краткие сведения о базах Web of Science, Scopus, Erih Plus

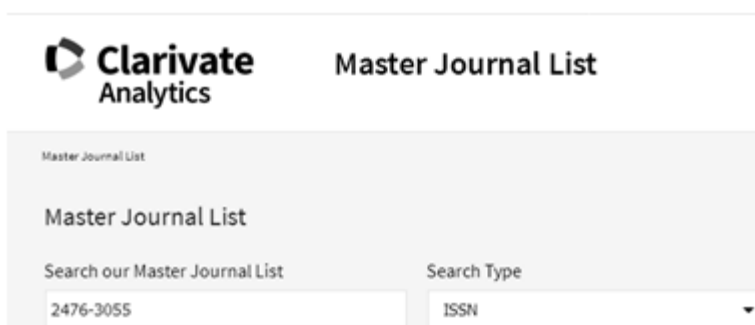
База данных Web of Science до недавнего времени разрабатывалась и принадлежала компании Thomson Reuters. Сейчас она управляется новой независимой компанией Clarivate Analytics. База Scopus принадлежит крупнейшему в мире научному издательству Elsevier, которое работает в партнерстве с лабораторией Scimago. База Erih Plus была создана Европейским научным фондом и сейчас принадлежит Норвежской службе социальных данных Norwegian Social Science Data Services.

Базы Web of Science и Scopus на текущий момент времени являются и наукометрическими, и библиографическими. База Erih Plus сейчас считается только библиографической и позиционируется в первую очередь как источник информации научного характера.

Подчеркнем, что сегодня прямой доступ в научные базы данных, как правило, либо невозможен, либо ограничен необходимостью регистрации и прочими сопутствующими

неудобствами (например, полноценный вход в базу Web of Science можно осуществить только с компьютера организации, и то, только в том случае, если она подключена к этой базе данных).

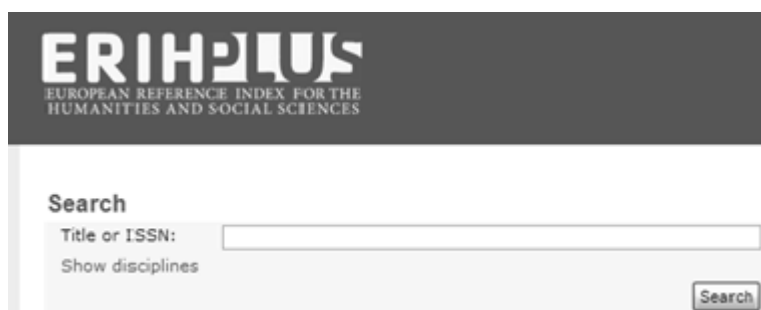
По этой причине, для того чтобы проверить, входит ли выбранный автором журнал в интересующую базу данных, лучше воспользоваться дополнительными программными ресурсами [Master Journal List, www; Scimago Journal & Country Rank, www; NSD ERIH PLUS, www], поставляемыми в свободном доступе либо самими организаторами баз данных, либо их партнерами. Подобные поисковые программы (рисунок 1), как правило, позволяют проверить наличие журнала в базе по его названию, либо по уникальному номеру периодического издания (ISSN). Он представляет собой восьмизначный код, состоящий из двух частей (по четыре знака, разделенных дефисом).



a)



б)



в)

Рисунок 1 - Фрагменты окон поисковых программ для баз данных:

a) – Master Journal List для Web of Science; *б)* – SJR для Scopus; *в)* – NDS для Erih Plus

Но, вместе с тем, эти вспомогательные ресурсы содержат лишь самую необходимую информацию: название, сведения о том, был ли включен журнал в базу данных, статус на текущий момент времени (включен\исключен), страну и город издательства, номер ISSN, краткое описание научных направлений журнала. В программе SJR для Scopus указывается также информация о ранге журнала и временном периоде вхождения журнала в базу.

Таким образом, подобные вспомогательные программы содержат набор только самых необходимых сведений об интересующих изданиях, индексируемых в базе данных. Для более полной информации следует работать непосредственно с самими базами данных в режиме полного доступа.

Критерии оценки периодических изданий, импакт-фактор и ранг журнала

Одной из важнейших функций наукометрической базы данных является сравнительная оценка периодических изданий. Например, основным численным показателем рейтинга журнала в базе Web of Science является, так называемый, импакт-фактор (IF) (таблица 1). Расчет импакт-фактора основан на трехлетнем периоде [Импакт-фактор, www]. Импакт-фактор журнала в текущем году вычисляется по предыдущим трем годам и рассчитывается как дробь A/B , в которой A - число цитирований, а B - число статей за прошедшие три года.

Таблица 1 - Показатели эффективности научных периодических изданий в различных базах данных

База данных	Наименование показателя
Web of Science	Импакт фактор IF
Scopus	Ранг журнала SJR
Erih Plus	-

Для оценки авторитетности журнала в Scopus используется аналогичный критерий [SJR, www], называемый рангом журнала (SJR). В базе Erih Plus подобный показатель отсутствует, поскольку сейчас эта база не является наукометрической.

Таким образом, делаем вывод, что чем больше импакт-фактор (ранг) журнала, тем выше цитируемость и рейтинг журнала, тем больше у потенциального автора научной статьи быть процитированным после публикации своей работы. Однако, с другой стороны, следует отметить, что с повышением рейтинга издания повышаются и требования к оформлению и содержанию научных работ, предоставляемых авторами в редакцию. Кроме того, вне зависимости от рейтинга издания, все присылаемые в редакцию рукописи проходят обязательную процедуру рецензирования.

Критерии оценки ученых и организаций, индекс цитирования и индекс Хирша

Наиболее простым и интуитивно понятным критерием оценки ученых и организаций является индекс цитирования [Индекс цитирования научных статей, www]. В простейшем случае он представляет собой число, равное общему количеству всех цитирований публикаций автора. Рассмотрим особенности этого критерия.

Предположим, что количество публикаций автора нарастает линейно, то есть каждый год автор издает примерно одинаковое количество трудов (это хорошо согласуется с

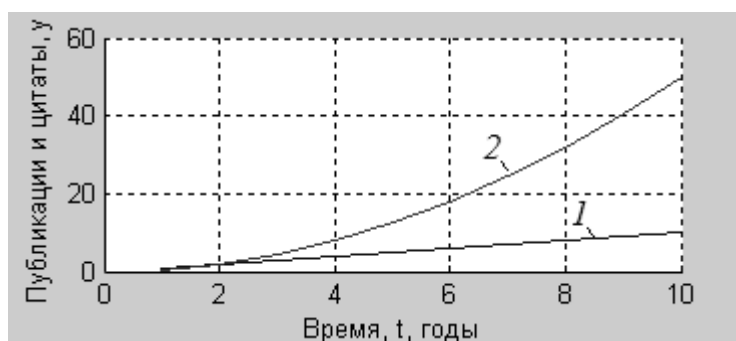
действительностью). Однако, с точки зрения сравнительного анализа, очевидно, что в этом случае индекс цитирования так же линейно нарастать не будет, поскольку каждая новая публикация открывает новую дополнительную возможность для цитирования, а прежние публикации также остаются вовлеченными в процесс цитирования наряду с новой работой.

Иными словами, чем больше публикаций имеет автор, тем выше темп цитирования его научных работ. Математически это можно выразить так:

$$N_1(t) = k_1 t, \quad N_2(t) = k_2 \int_{T_1}^{T_2} (k_1 t) dt, \quad (1)$$

где: $N_1(t)$ – функция количества публикаций автора; k_1 – коэффициент публикационной активности автора; t – время; $N_2(t)$ – функция общего количества цитирований публикаций автора; k_2 – коэффициент цитируемости; T_1, T_2 – границы временного интервала наблюдения.

Поясним данную особенность индекса цитирования. Пусть, к примеру, автор издает по одной индексируемой публикации в год (коэффициент k_1 равен единице), а каждая изданная статья автора цитируется также один раз в год (коэффициент k_2 также равен единице), начиная с момента публикации (рисунок 2). Тогда за первый год будет один раз процитирована только первая статья автора. За второй год будет еще раз процитирована первая статья и в первый раз процитирована вторая статья. За третий год еще по разу будут процитированы первая и вторая статьи, а также в первый раз процитирована третья и так далее (то есть, по смыслу, приходим к интегральной зависимости).



Пояснение к принципу прироста публикаций и индекса цитируемости автора с обозначениями: 1 – график прироста публикаций; 2 – график прироста индекса цитирования

Рисунок 2 – Прирост публикаций и цитирований

В результате за 10 лет получаем 10 статей, но уже 50 цитирований. То есть, прирост количества статей автора происходит линейно, а при формировании индекса цитирования прослеживается нелинейная параболическая зависимость. Иными словами, темп публикации статей автора постоянен, а темп цитируемости нарастает с каждой добавляющейся статьей. С точки зрения математики этот темп характеризует первая производная записанных функций (y у $N_1(t)$ она является константой, а у $N_2(t)$ линейно-нарастающей функцией, зависящей от t).

В результате можно получить парадоксальную ситуацию, когда два сравниваемых ученых имеют сопоставимые количества публикаций, но несоизмеримо различающиеся индексы цитирования. Причем эта разница может являться следствием объективных причин (один из

ученых действительно востребованней), а может являться следствием рассмотренной особенности роста индекса цитирования.

То есть здесь речь идет о том, что количество не всегда означает качество. Кроме того, одни предметные области могут быть востребованнее других. Тема исследования может быть широкой или узкоспециализированной. В научном мире могут иметь место такие аспекты как самоцитируемость, цитируемость соавторами и так далее.

Вернемся к критериям. Еще одним критерием сравнительной оценки ученых и организаций в наукометрических базах данных выступает индекс Хирша [Индекс Хирша, [www](#)]. Данный индекс сглаживает нелинейность, присущую графику обычного индекса цитирования. То есть, в первом приближении можно сказать, что темп роста этого индекса уже не так сильно зависит от количества публикаций ученого, накопленных на текущий момент времени.

При вычислении индекса Хирша действует правило: ученый имеет индекс h , если h из общего числа его N статей цитируются как минимум h раз каждая, в то время как оставшиеся $N - h$ статей цитируются не более чем h раз каждая. Таким образом, для оценки данной величины свои публикации необходимо расположить на оси абсцисс в порядке убывания количества цитирований, на оси ординат обозначить количества цитат, и далее искать число h , двигаясь слева направо по оси абсцисс.

Приведем конкретные примеры. Индекс Хирша ученого, имеющего 3 публикации, процитированные каждая по 3 раза имеет индекс Хирша 3. Индекс Хирша ученого, имеющего 5 публикаций, две из которых процитированы по 4 раза, а остальные не цитировались, имеет индекс Хирша 2. Индекс Хирша ученого, имеющего 10 публикаций, каждая из которых цитировалась по 1 разу, равен 1.

Таким образом, чем больше становится индекс Хирша, тем больше требуется цитирований, чтобы получить следующий пункт. По этой причине индекс Хирша ученого обычно лежит в пределах нескольких десятков. Например, в области физики состоявшимся считается тот ученый, который имеет индекс Хирша не менее 10. Для сравнения, Нобелевские лауреаты имеют индекс Хирша порядка 60 (то есть, имеются 60 публикаций, каждая из которых процитирована по 60 раз, что в итоге дает 3600 цитирований).

Возвращаясь к приведенным ранее графикам, отмечаем, что поскольку нелинейность роста, присущая простому индексу цитирования, в данном случае сглажена, то индекс Хирша можно считать более объективным критерием, чем простой индекс цитирования. Однако, с другой стороны, следует заметить, что и этот индекс не является «свободным» от аспектов, связанных с самоцитированием, цитированием соавторами, различий в отраслях науки, специфике исследований и прочих факторов.

Заключение

В целом отметим, что с точки зрения образования все описанные базы данных являются наиболее полными и исчерпывающими ресурсами для поиска информации. Поэтому в плоскости учебного процесса их целесообразно рассматривать, прежде всего, как библиографические и реферативные источники информации, способствующие подготовке лекций, курсовых проектов, диссертаций и так далее.

Применительно к науке, рассмотренные базы (кроме Erih Plus) следует позиционировать как наукометрические базы данных, которые не только регистрируют, но и позволяют проводить численную оценку рейтингов научных журналов, а также сравнивать показатели

научных организаций и отдельно взятых ученых.

При этом потенциальная польза от использования наукометрических баз данных в целях развития науки и образования сомнений не вызывает. Сама попытка численно выразить рейтинг периодических изданий, эффективность работы научных организаций, ученых и сотрудников образовательных учреждений уже сама по себе представляется интересной, а главное востребованной временем.

Однако, с другой стороны, очевидно также, что данная оценка принесет наибольшую практическую пользу обществу, науке и образованию только тогда, когда она будет максимально объективной. В этом смысле для повышения уровня качества и объективности оценки следует уделить пристальное внимание таким аспектам наукометрических баз данных как: актуальность темы исследования, различие отраслей науки, специфичность предметной области, самоцитирование, цитирование соавторами, время публикационного периода и так далее.

Библиография

1. Импакт-фактор. URL: <https://ru.wikipedia.org/wiki/Импакт-фактор>
2. Индекс Хирша. URL: https://ru.wikipedia.org/wiki/Индекс_Хирша/
3. Индекс цитирования научных статей. URL: https://ru.wikipedia.org/wiki/Индекс_цитирования_научных_статей
4. ERIH PLUS. URL: https://en.wikipedia.org/wiki/ERIH_PLUS
5. Master Journal List. URL: <http://mjl.clarivate.com/>
6. NSD ERIH PLUS European Reference Index for the Humanities and Social Sciences. URL: <https://dbh.nsd.uib.no/publiseringskanaler/erihplus/>
7. Scimago Journal & Country Rank. URL: <https://www.scimagojr.com/>
8. Scopus. URL: <https://ru.wikipedia.org/wiki/Scopus>
9. SJR, Scimago Journal & Country Rank. URL: <https://www.scimagojr.com/journalrank.php>
10. Web of Science. URL: https://ru.wikipedia.org/wiki/Web_of_Science

The role of scientometric and bibliographic databases in the fields of science and education

Dmitrii S. Kalistratov

PhD in Technical Science, Applicant, Assistant,
Tula State University,
300012, 92, Lenina ave., Tula, Russian Federation;
e-mail: Kalistratow@list.ru

Abstract

The role of scientometric and bibliographic databases in the fields of science and higher vocational education is analyzed, the concepts of metrics, metric space and scientometric database are revealed, brief information about the most common scientific databases Web of Science, Scopus and Erih Plus is provided, their similarities and differences are indicated, references to the Internet resources containing data on inclusion of periodicals in the specified databases are given, from the point of view of objectivity and usefulness for science and education the main criteria of comparative evaluation of periodicals, individual scientists, as well as scientific and educational institutions are investigated, in particular, the concepts relating to the rating of scientific publications, studied the

concept of impact factor and the rank of the periodical, on specific examples the features of the growth of the citation index and the h-index of a single scientist are demonstrated, a comparative assessment of these criteria is given, their advantages and disadvantages are indicated, thanks to analysis positive and negative aspects of the use of scientometric and bibliographic databases in the fields of science and education are revealed, the role of scientometric databases for science, education and society as a whole is also indicated here.

For citation

Kalistratov D.S. (2019) Rol' naukometricheskikh i bibliograficheskikh baz dannykh v sferakh nauki i obrazovaniya [The role of scientometric and bibliographic databases in the fields of science and education]. *Pedagogicheskii zhurnal* [Pedagogical Journal], 9 (1A), pp. 87-94. DOI: 10.34670/AR.2019.44.1.032

Keywords

Science, education, pedagogy, scientometric database, advantages, disadvantages, impact factor, publication rank, citation index, h-index.

References

1. *ERIH PLUS*. Available at: https://en.wikipedia.org/wiki/ERIH_PLUS [Accessed 12/12/2018]
2. *Impakt-faktor* [Impact factor]. Available at: <https://ru.wikipedia.org/wiki/Impakt-faktor> [Accessed 12/12/2018]
3. *Indeks Khirsha* [The Hirsch Index]. Available at: https://ru.wikipedia.org/wiki/Indeks_Khirsha/ [Accessed 12/12/2018]
4. *Indeks tsitirovaniya nauchnykh statei* [The citation index of scientific articles]. Available at: https://ru.wikipedia.org/wiki/Indeks_tsitirovaniya_nauchnykh_statei [Accessed 12/12/2018]
5. *Master Journal List*. Available at: <http://mjl.clarivate.com/> [Accessed 12/12/2018]
6. *NSD ERIH PLUS European Reference Index for the Humanities and Social Sciences*. Available at: <https://dbh.nsd.uib.no/publiseringsskanaler/erihplus/> [Accessed 12/12/2018]
7. *Scimago Journal & Country Rank*. Available at: <https://www.scimagojr.com/> [Accessed 12/12/2018]
8. *Scopus*. Available at: <https://ru.wikipedia.org/wiki/Scopus> [Accessed 12/12/2018]
9. *SJR, Scimago Journal & Country Rank*. Available at: <https://www.scimagojr.com/journalrank.php> [Accessed 12/12/2018]
10. *Web of Science*. Available at: https://ru.wikipedia.org/wiki/Web_of_Science [Accessed 12/12/2018]