

НАУКОМЕТРИЯ. БИБЛИОМЕТРИЯ

УДК [001.83:01]-047.44

<https://doi.org/10.33186/1027-3689-2023-2-37-58>

Вызовы для развития наукометрических исследований

А. Е. Гуськов¹, Я. Л. Шрайберг²

¹*Российский научно-исследовательский институт экономики, политики и права в научно-технической сфере, Москва, Российская Федерация, guskov.andrey@gmail.com*

²*ГПНТБ России, Москва, Российская Федерация; Московский государственный лингвистический университет, Москва, Российская Федерация, gpntb@gpntb.ru*

Аннотация. В статье рассматриваются основные вызовы – сложные проблемы, препятствующие развитию наукометрических исследований и научных коммуникаций. Для их решения требуются большие усилия и профессиональная смелость. В первую очередь, к таким вызовам относятся необходимость открытого доступа к наукометрическим данным, а также повышение их качества и полноты, включая сведения об авторах, аффилиациях, цитированиях и иную метаинформацию. Обоснована необходимость массового применения технологий идентификации научно-информационных объектов (публикаций, исследователей, организаций, проектов и др.), что позволит значительно снизить количество библиографических ошибок.

При проектировании наукометрических изысканий границы объектов исследований и инструменты анализа должны определяться их целями, а не ограничениями библиометрических баз данных. Одним из ключевых инструментов является классификация научных публикаций. Его развитие обусловлено появлением новых и низким качеством существующих в библиометрических базах данных классификаторов, их различиями, а также изменчивостью структуры науки. Наконец, необходимо строго следить за корректностью применения методов наукометрического анализа и интерпретации его результатов, особенно при оценке научной результативности.

Преодоление этих вызовов позволит реализовать эффективный мониторинг научной деятельности, который будет основываться уже не на формах ежегодного статистического наблюдения, а на процедурах оперативного сбора, обработки и анализа потоков научной информации. Такой переход позволит

значительно улучшить характеристики мониторинга, расширить спектр решаемых задач, выявлять системные изменения в сфере исследований, своевременно реагировать на дисбалансы в её развитии и увеличить эффективность решений, принимаемых в управлении научной деятельностью.

Ключевые слова: наукометрия, библиометрия, Web of Science, Scopus, DOI, ORCID, классификация, метаданные

Для цитирования: Гуськов А. Е., Шрайберг Я. Л. Вызовы для развития наукометрических исследований // Научные и технические библиотеки. 2023. № 2. С. 37–58. <https://doi.org/10.33186/1027-3689-2023-2-37-58>

SCIENTOMETRICS. BIBLIOMETRICS

UDC [001.83:01]-047.44

<https://doi.org/10.33186/1027-3689-2023-2-37-58>

Challenges to develop scientometric studies

Andrey E. Guskov¹ and Yakov L. Shrayberg²

¹*Russian Research Institute of Economics, Politics and Law in Science and Technology, Moscow, Russian Federation, guskov.andrey@gmail.com*

²*Russian National Public Library for Science and Technology, Moscow, Russian Federation; Moscow State Linguistic University, Moscow, Russian Federation, gpntb@gpntb.ru*

Abstract. The authors examine the key problems inhibiting scientometric studies and scientific communications. These challenges call for significant efforts and professional courage. Firstly, this is the need for open access to scientometric data and improvement of their quality and comprehensiveness, including author data, affiliations, citations and meta information. The authors emphasize the necessity for large-scale introduction of technologies for identifying objects of science information (i. e. publications, researchers, organizations, projects, etc.), which would enable to decrease significantly the number of bibliographic mistakes.

When projecting scientometric studies, the edge of objects and analysis instruments have to be defined by the goals rather than by bibliometric database limitations. Indexing of scientific publications is among the key instruments. Its advancement is determined by emerging and low-quality classifications of bibliometric databases, their differences, and changing science structure. Finally, the propriety of scientometric methods and results interpretation, in particular that of scientometric performance assessment, have to be controlled. Meeting these challenges will enable to provide efficient monitoring of scientific activity based on operative collection, processing and analysis of scientific information flows rather than on annual statistical surveys. This transfer would improve monitoring significantly and expand the spectrum of solutions; it would also enable to reveal system changes in research, to respond to disparities in development, and to make the solutions in science management more efficient.

Keywords: scientometrics, bibliometrics, Web of Science, Scopus, DOI, ORCID, classification, metadata

Cite: Guskov A. E., Shrayberg Y. L. Challenges to develop scientometric studies // Scientific and Technical Libraries. 2023. No. 2. P. 37–58. <https://doi.org/10.33186/1027-3689-2023-2-37-58>

Развитие методов наукометрического анализа сегодня хорошо согласуется с повсеместным применением технологий больших данных и цифровизацией различных областей деятельности. Внедрение информационных технологий зачастую сопровождается новыми вызовами – сложными проблемами, для решения которых требуются большие усилия и профессиональная смелость. Преимущественно они связаны с доступностью, качеством и верифицируемостью первичных и вторичных данных, а также с необходимостью реализации новых алгоритмов их сбора, гармонизации и анализа.

В этой работе рассматриваются основные вызовы, препятствующие развитию наукометрических исследований. Они могут заключаться в недостаточном качестве результатов исследований и высоких статистических погрешностях, невозможности выполнения некоторых видов анализа или его масштабирования на более крупные объекты (переход от анализа отдельных организаций или групп к регионам или странам),

а также в некорректном применении наукометрических данных. Этот список проблем не следует считать исчерпывающим, но это основные задачи, с которыми сталкиваются исследователи в нашей области (особенно если они приходят не вполне подготовленными). Тем, кто применяет или собирается применять наукометрические исследования в библиотечно-информационной практике, необходимо понимать, что такое научная коммуникация в библиотечно-информационной сфере, а также иметь представление о классификационных системах в интернете и о том, как проходит классификация научно-технического, в частности гуманитарного знания, с помощью таблиц ББК и УДК [1–3].

1. Доступность данных для наукометрического анализа

Первым препятствием на пути развития наукометрических исследований является ограничение доступа к исходным данным о публикациях, исследователях, организациях, проектах и др.

Основной источник данных для наукометрических исследований – библиографические базы данных (ББД), которые также часто называют наукометрическими или базами цитирования. Изначально эти системы создавались с целью формирования каталогов научных публикаций, отвечающих определённым характеристикам (например, по общему тематическому признаку или на основании списка верифицированных источников). Однако, благодаря одному из трендов современной информатизации – связыванию данных из различных источников («from cataloguing to catalinking» [4]), в этих ББД были выделены и идентифицированы сопутствующие вторичные сущности: издательства, журналы, конференции, авторы, места их работы, источники финансирования и др. Это позволило перейти от анализа массива научных публикаций к более комплексному анализу результатов и структуры научной деятельности.

В мировой практике наукометрических исследований чаще всего используются две ББД: Web of Science (владелец – Clarivate Analytics) и Scopus (Elsevier). Реже упоминается Google Scholar, имеющая больший охват, но худшее качество данных и инструментов для наукометрического анализа.

Впрочем, сегодня и Web of Science, и Scopus закрыли доступ к своим базам данных для российских пользователей, и это – уже новая проблема, которую необходимо срочно решать. Владельцы вышеуказанных ресурсов предоставляли свои программные интерфейсы широкому кругу пользователей, в том числе безвозмездно при наличии подписки на основную базу данных.

Формирование общедоступного массива публикаций стало возможным с появлением в 2000 г. технологии DOI (Digital object identifier) [2] и реализующей её системы Crossref, в которой регистрируются не только идентификаторы, но и метаданные. К этой системе присоединились далеко не все научные издатели, однако их количество постоянно растёт и уже вполне сопоставимо с Web of Science и Scopus.

После 2015 г. появилось несколько новых библиографических систем, таких как Dimensions, Lens, OpenAlex и Semantic Scholar, которые также имеют большое (хотя и заметно меньшее) покрытие научных источников и API для поисковых и аналитических запросов [6, 7].

Таким образом, концептуальные вопросы обеспечения доступности библиографических данных для наукометрического анализа могут считаться решёнными. При этом требуемое для наукометрических исследований качество открытых библиографических данных пока не обеспечивается, и это является отдельным вызовом. Отметим, что методические подходы, прежде всего в области оценки цитируемости научных статей российских журналов, разрабатываются [8, 9].

Объектами наукометрического анализа являются не только публикации, но и исследователи, университеты, научные проекты, конференции и др. Использование реестров и систем идентификации (например, ROR для организаций, ORCID для исследователей [5]) пока не столь масштабно, что затрудняет их применение для задач наукометрического анализа.

2. Качество данных для наукометрического анализа

Наукометрические данные должны содержать корректные и полные сведения об описываемом объекте.

Причины низкого качества данных разнообразны: различие форматов метаданных, перевод значений полей с одного языка на другой, использование нелатинского алфавита, формул и названий химических

соединений, игнорирование необязательных полей при вводе, неточности при вводе данных (человеческий фактор). Ошибки при создании библиографических записей или их обновлении могут допускать авторы публикаций, редакторы и издатели, научные агрегаторы.

Проведённое сравнение показало заметные различия полноты и качества метаданных в разных системах [6]. Наилучшие показатели продемонстрировали старожилы рынка научной информации Web of Science и Scopus, которые имеют свои собственные процедуры сбора и «улучшения» данных.

Следует отметить, что издатели не всегда одобряют открытое пространство своих метаданных, так как считают, что метаданные обладают коммерческой ценностью и являются конкурентным преимуществом. Для преодоления этих барьеров появилось две инициативы – Initiative for Open Citations (i4oc.org) и Initiative for Open Abstracts (i4oa.org), предоставляющих Открытый доступ к аннотациям статей и спискам цитирований через систему Crossref. Ко второй инициативе до сих пор не присоединились крупнейшие научные издатели Elsevier, Springer Nature, Wiley и Taylor & Francis. Как следствие, в 2021 г. в открытой системе Crossref были доступны аннотации лишь 30% статей.

Таким образом, несмотря на относительно высокую доступность наукометрических данных, их качество остаётся очень неоднородным. Необходимы систематические усилия научного и издательского сообществ по его улучшению.

3. Идентификация научно-информационных объектов

В наукометрических базах данных должна быть возможность идентифицировать связанные научно-информационные объекты: публикации, исследователей, организации, проекты, научные фонды, журналы, конференции и др.

Идентификация сущностей, являющаяся частью процедур повышения качества метаданных, представляет собой отдельный вызов, так как с ней связан класс специализированных систем, обеспечивающих идентификацию научно-информационных объектов. К таким системам относятся DOI (публикации) [10], ORCID [11], ResearcherID, SPIN-код (исследователи), GRID, ROR, OpenRIRO (организации), ISBN (книги), ISSN (журналы) и др. Преодоление этого вызова необходимо для совместно-

го наукометрического анализа двух и более научно-информационных объектов (например, публикации и авторы, проекты и организации), а также для интеграции наукометрических данных из различных источников.

Применение систем идентификации достаточно широко, но не тотально. При их отсутствии используются разнообразные алгоритмы идентификации по именам, названиям и другим факторам с достаточно большой погрешностью, приводящей к дублированию научных объектов. Проведённое в 2019 г. исследование [12] выявило, что 76% организаций и 24% авторов научных публикаций имеют профили-дубли в Scopus. Дубли профилей организаций появляются по следующим причинам: различия при транслитерации и переводе названий организаций, опечатки и ошибки в названиях, отсутствие части названия, указание разных адресов, неверное указание принадлежности к тому или иному региональному отделению РАН. Дубли профилей авторов также возникают из-за различий при транслитерации имён, фамилий и возможных опечаток. Типичная причина – неверное написание фамилии, имени или отчества, например, к фамилии добавляется символ из обозначения организации или отчества; фамилия и имя могут быть поменяны местами.

Таким образом, при наукометрическом анализе результативности научной деятельности необходимо учитывать погрешность, вызванную некорректной идентификацией и связыванием авторов, организаций и публикаций. Применение систем идентификации научно-информационных сущностей значительно уменьшает последствия таких ошибок.

4. Расширение охвата наукометрических исследований

Границы объектов наукометрических исследований должны определяться целями исследований, а не ограничениями библиометрических баз данных.

Как уже отмечалось, большинство наукометрических исследований выполняется на основе данных Web of Science или Scopus, в каждой из которых существует строгая процедура отбора наиболее авторитетных источников для включения в базу. Однако они составляют лишь небольшой срез, верхушку айсберга, большая часть которого не видна. Несмотря на то, что в научном сообществе не вошедшие в базу

источники пользуются меньшим авторитетом, в них также могут содержаться значимые результаты исследований (хотя их концентрация обычно ниже).

В качестве примера можно привести российскую научную периодику: в Web of Science и Scopus индексируются 300–400 журналов, в списке Высшей аттестационной комиссии (ВАК) присутствует более 2 тыс. журналов, а всего в Российском индексе научного цитирования (РИНЦ) зарегистрировано около 6 тыс. периодических отечественных изданий. Таким образом, любой анализ на основе только международных баз данных будет охватывать не более 10% российских изданий, которые позиционируются как научные. Этого может быть достаточно для анализа общих тенденций в научной сфере или отдельных дисциплинах, однако полнота таких исследований всегда будет под вопросом.

Особенно существенно это влияет на общественно-гуманитарные дисциплины с выраженной локализованностью тематики в рамках территории, языка, этноса и т. п. В основном они представляют интерес для региональных групп учёных и гораздо реже для зарубежных исследователей. Из-за этого соответствующим научным изданиям сложно попасть в индексы Web of Science и Scopus, а сами тематики могут быть недостаточно представлены в наукометрических исследованиях.

Формально отличить настоящую науку от симулирующей или «туземной» невозможно [13]. Поэтому при любом расширении охвата появляется риск, что среди источников окажется множество откровенно слабых с научной точки зрения журналов, публикации в которых не содержат принципиально новых результатов. Качество наукометрических исследований, основой которых будет массив публикаций совершенно разного уровня, понизится. Поэтому при расширении охвата следует применять методы квалифицированного отбора источников, которые обеспечат не только полноту массива наукометрических данных, но и необходимый уровень качества.

5. Разработка инструментов наукометрического анализа

Усложнение задач наукометрического анализа требует усложнения инструментов для их выполнения.

Для проведения наукометрических исследований в Web of Science и Scopus были разработаны соответствующие специализированные

инструменты – InCites и SciVal. Они позволяют строить, фильтровать, агрегировать данные о публикациях и выполнять аналитические срезы по различным параметрам, используя разные формы: графики временных рядов, круговые диаграммы, карты и др. Оценку международной научной активности в сфере библиотековедческих исследований за 10 лет (2011–2020) на основе платформы SciVal проанализировала Т. В. Еременко [14]. Другими примерами аналитических инструментов являются свободно распространяемые программы VOSViewer [15] и CiteSpace [16], которые на основе заданной выборки публикаций строят графы, кластеры и тепловые карты научного ландшафта. На сегодняшний день это наиболее популярные инструменты наукометрического анализа.

Пользоваться ББД и сопутствующими аналитическими системами может человек со средними навыками компьютерной грамотности и без особой подготовки (не считая небольшого курса по наукометрическим показателям). Однако основное ограничение упомянутых выше инструментов состоит в невозможности проведения анализа, который в явном виде не включён в их функционал. Несмотря на разнообразие функционала, невозможно далеко выйти за рамки простого аналитического процесса «фильтр – агрегация – визуализация». Этого недостаточно для решения многих задач, в которых требуется прямое и итерационное оперирование исходным массивом данных о публикациях, авторах, организациях, научных проектах и других сущностях с применением математического аппарата. Во многих оригинальных наукометрических инструментах используются стандартные пакеты программного обеспечения для обработки данных на языках Python, R, Java и др., но они требуют привлечения высококвалифицированного ИТ-специалиста с математической подготовкой.

Когда набор оригинальных наукометрических исследований проходит апробацию и становится типовым сценарием наукометрического анализа, возникает потребность в специальном инструменте для его реализации (или в новой функции у имеющегося). Количество таких сценариев растёт, они становятся более сложными и составными. Возникает необходимость во взаимном комплексировании аналитических инструментов, взаимодействии с различными источниками данных, их интеграции и гармонизации.

6. Тематическая классификация научных текстов

Каждая научная публикация должна быть отнесена к одному или нескольким рубрикам тематического классификатора. При этом разным задачам могут соответствовать разные классификаторы, структура которых может со временем изменяться.

Актуальность автоматической классификации научных публикаций по тематикам выросла в связи со значительным и постоянным ростом их числа. По каждой научной дисциплине ежегодно появляются сотни, тысячи и десятки тысяч статей. Проблема усугубляется существованием нескольких параллельных систем классификации, плохо согласующихся друг с другом. Например, Web of Science, Scopus и РИНЦ используют разные тематические рубрикаторы. При попытке слияния массивов из этих систем возникают сложности с их общей классификацией.

Следующая проблема заключается в использовании так называемой журнальной классификации, когда всем статьям присваиваются имеющиеся у журнала рубрики. Такой подход работает для монотематических журналов, в остальных случаях он становится менее корректным, а в мультидисциплинарных журналах и вовсе бессмысленным.

Ещё одна проблема обусловлена изменчивостью структуры науки, появлением новых дисциплин (нанотехнологии, большие данные, машинное обучение), что приводит к необходимости периодического пересмотра структуры классификаторов, а значит и к повторной классификации всего массива публикаций.

Аналогичная задача возникает и при введении новых классификаторов. Например, в России определены семь приоритетных направлений научно-технологического развития, которые были закреплены в целях Национального проекта «Наука». Возникла задача классификации результатов научной деятельности по этим направлениям. Предложенный способ решения (разработка «конвертора», который относит тематические категории журналов Web of Science к одному из приоритетов) можно назвать удовлетворительным лишь с большой натяжкой – приоритеты сформулированы в виде глобальных задач, для решения которых должны применяться результаты из разных областей наук.

Наконец, являются востребованными классификаторы разных уровней детализации для различных целей. В одних случаях для задач стратегического управления сферой R&D требуется верхнеуровневая маркировка большого потока публикаций на несколько разделов (как в примере выше), в других – необходимо выделить небольшие кластеры близких по тематике публикаций (исследовательские топики), причём количество кластеров может исчисляться десятками и сотнями тысяч.

Таким образом, для проведения широкого спектра наукометрических исследований необходимо решение задачи автоматической тематической классификации больших массивов научных публикаций. Решением данной задачи является создание единой сети связей классификаций, базовой классификацией которой является Государственный рубрикатор научно-технической информации (ГРНТИ). Разработка системы взаимосвязанных классификаций обеспечит информационную совместимость, автоматическую тематическую систематизацию и интеграцию потоков научных публикаций, входящих в состав информационных подпространств, систематизированных различными классификациями.

7. Применение результатов наукометрического анализа

Необходимо строго следить за корректностью применения методов наукометрического анализа и интерпретацией его результатов, особенно при оценке научной результативности.

Научные исследования являются сложным для управления объектом с большим количеством неявных внутренних связей, инерционностью и крайне неочевидной зависимостью результата от входных условий. Наукометрический подход для некоторых административных задач частично упрощает эту сложность, задавая систему мер и весов. При этом, чем значительнее такое упрощение и чем проще предлагаемая наукометрическая модель, тем хуже она отражает реальную ситуацию, создавая неверное представление о ней и предпосылки для ошибочных решений.

Рассмотрим пример, когда найм сотрудника на научную ставку в организации основан только на индексе Хирша. С одной стороны, это избавляет от необходимости вникать в детали резюме соискателей

и минимизирует субъективность решения: выбор из двух кандидатов с индексом Хирша 10 и 12, очевидно, будет сделан в пользу последнего. При этом может оказаться, что первый соискатель в два раза моложе, и, вероятно, имеет более высокий научный потенциал. Однако слишком простая система оценки не позволила отследить этот фактор.

Второй пример касается стратегического целеполагания. В национальном проекте «Наука и университеты» одним из ключевых является показатель «Место Российской Федерации по объёму научных исследований и разработок». Согласно паспорту проекта наша страна должна подняться с 10-го места в 2020 г. на 8-е место в 2024 г. В методике Минобрнауки, утверждённой распоряжением № 421-Р от 02.11.2021, расчёт определялся по взвешенной сумме из пяти параметров, одним из которых является показатель «Место Российской Федерации по удельному весу в общем числе изданий, индексируемых в международных базах данных» (подразумеваются Web of Science и Scopus). При этом методика не конкретизирует, какие виды статей (или, возможно, публикаций) учитываются при расчёте.

В подобных случаях применяются несколько вариантов расчётов: все публикации, только статьи в научных журналах (AR), статьи и обзоры в научных журналах (AR + RE). Результат в значительной степени зависит от методики: в первом случае Россия находится на 8-м месте, а во втором – лишь на 12-м.

Место России среди других стран по общему числу всех публикаций, статей в журналах и статей в трудах конференций, проиндексированных в БД Scopus в 2022 г.

Место	Все публикации		Статьи в журналах (AR)		Статьи в трудах конференций (CP)	
	Страна	Число	Страна	Число	Страна	Число
1	China	772 112	China	607 315	China	110 949
2	USA	709 033	USA	485 592	USA	78 354
3	UK	227 251	UK	155 579	India	44 684
4	India	214 341	India	137 044	Russia	38 848
5	Germany	196 479	Germany	134 511	Germany	27 355
6	Italy	146 903	Japan	103 098	Indonesia	24 428
7	Japan	138 710	Italy	97 095	Japan	19 490

Место	Все публикации		Статьи в журналах (AR)		Статьи в трудах конференций (CP)	
8	Russia	128 449	France	88 305	UK	18 702
9	France	126 112	Spain	86 862	Italy	16 329
10	Canada	122 172	Canada	86 246	France	14 684
11	Australia	117 044	Australia	84 236	Canada	12 065
12	Spain	114 447	Russia	79 546	Australia	8 600

Причина расхождений – аналогичный рейтинг по количеству публикаций в трудах конференций, проиндексированных в Scopus, в котором Россия в 2020 г. вышла на четвёртую позицию. Оказывается, что доля таких публикаций у российских учёных уже более чем в два раза превышает среднемировой уровень (он составляет 15–16% от всех публикаций). Известно, что требования к публикациям в трудах многих конференций ниже, чем в научных журналах: упрощённая процедура рецензирования отражается на их качестве. Феномен российского «конференционного взрыва» изучался И. А. Стерлиговым, который пришёл к выводу о том, что «наблюдаемые колоссальные различия между российской и иностранной практикой, между вузами и РАН, наконец, между самими вузами заставляют в краткосрочной перспективе рекомендовать отказаться от учёта трудов конференций в официальных мониторингах наравне с журнальными статьями, собирать информацию об их количестве отдельно и применять только справочно» [17]. Схожие проблемы фиксировались и в других странах, например в Чехии [18] и Индонезии [19]. В каждом случае авторы связывают эти аномалии с неудачным дизайном системы национальной оценки, которая основывается на наукометрических показателях и оказывает чрезмерное «публикационное давление» на исследователей и администраторов науки, вынуждая их искать лёгкие пути для достижения показателей.

Мировое научное сообщество озабочено этой проблемой, о чём свидетельствует широкое распространение Декларации DORA (выступает против практики соотношения импакт-фактора журнала с вкладом конкретного учёного) [20] и Лейденского манифеста (оценка исследовательской деятельности, основанная на наукометрии) [21]. Однако

во многих странах эти проблемы всё ещё существуют, и Россия не исключение.

8. Изменение государственной системы мониторинга научной деятельности

Мониторинг научной деятельности должен основываться не на формах ежегодного статистического наблюдения, а на сборе, обработке и анализе потоков научной информации, что позволит значительно улучшить его характеристики, расширить спектр решаемых задач, выявить системные изменения в сфере исследований, своевременно реагировать на дисбалансы в её развитии и увеличить эффективность решений, принимаемых в управлении научной деятельностью.

Ключевым инструментом контроля за реализацией научной политики является государственный мониторинг состояния и результативности сферы науки, технологий и инноваций, осуществляемый Правительством РФ. Он основан на ежегодном сборе и анализе количественных показателей, отражающих уровень достижения результатов, в том числе в сопоставлении со значениями соответствующих показателей экономически развитых стран. Рассмотрим некоторые принципы организации этого мониторинга.

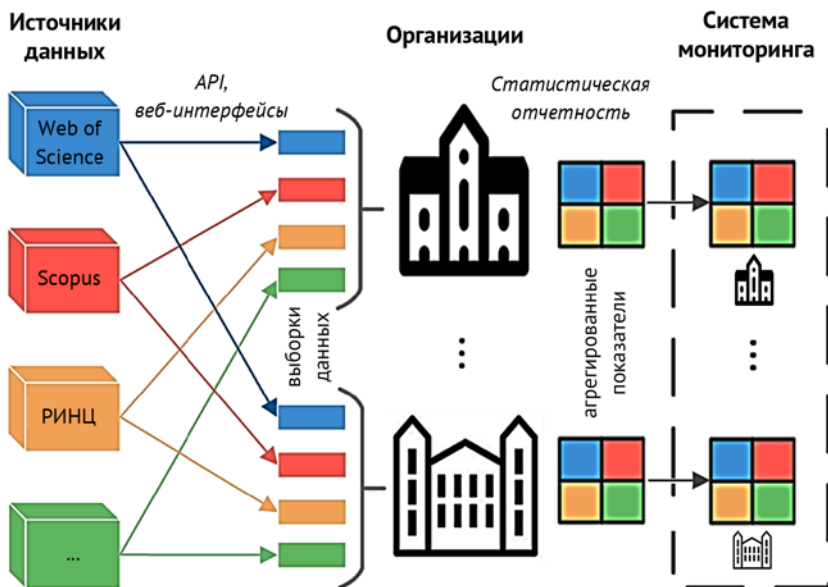
Основным источником количественной информации о результативности научных исследований и их кадровой обеспеченности до недавнего времени являлись отчётные формы федерального статистического наблюдения (ФСН), которые ежегодно заполняются научными организациями. Основная отчётная форма статистического мониторинга исследований «2-наука» была разработана в 1994 г. и с тех пор практически ежегодно изменялась для отслеживания новых тенденций (например, в 2008 г. в ней стали отдельно учитываться научные исследования и разработки, связанные с нанотехнологиями). При этом раздел 1 «Персонал, занятый научными исследованиями и разработками» содержит лишь сведения о количестве исследователей по областям наук, возрасту, полу и наличию учёной степени, а также о движении персонала, занятого научными исследованиями и разработками. Таким образом, содержание форм ФСН, которое определяется Росстатом, носит общий характер и не предоставляет достаточно сведений о научной деятельности, её результативности и кадровой обеспеченности.

В 2014 г. Федеральной службой по надзору в сфере образования и науки (Рособрнадзор) была запущена Федеральная система мониторинга научных организаций, выполняющих научно-исследовательские, опытно-конструкторские и технологические работы (ФСМНО), содержащая «сведения об оценке и о мониторинге результативности деятельности научных организаций, выполняющих научно-исследовательские, опытно-конструкторские и технологические работы гражданского назначения». Научные организации и вузы ежегодно предоставляли ей статистические сведения, состав которых был заметно шире, чем в формах ФСН. Для ряда показателей была введена детализация по научным дисциплинам на основе справочника кодов классификации отраслей наук Организации экономического сотрудничества и развития (ОЭСР – OECD). Важным нововведением стала необходимость подтверждения указанных организацией значений показателей путём предъявления скриншотов, поисковых запросов или идентификаторов.

Однако в целом сбор данных в ФСМНО также основывается на статистическом подходе (см. рис.), когда объектом сбора является агрегированный показатель (как правило, суммарное, среднее или нормированное значение), а исходные (или первичные) данные для его расчёта остаются за скобками. Это даёт возможность решить априори поставленные задачи, для которых и были разработаны эти показатели, но не позволяет провести более глубокий анализ данных.

Важно сформировать и проанализировать объективную картину происходящего, основанную на исходных данных, так как статистические подходы всё чаще не отражают реального положения вещей. Организации, участвующие в сборе данных, привносят в информационный поток плохо контролируемый «разрыв», который сопровождается ошибками разного масштаба и природы, в том числе – обусловленными человеческим фактором. Выше было показано, что сведения государственного статистического учёта имеют низкую скорость обновления, невысокий уровень достоверности. Они предназначены для решения заранее определённых задач; не вполне подходят для глубоких аналитических исследований (особенно прогнозного характера), а также для создания систем поддержки принятия решений [22]. Это означает, что для формирования «эффективной современной системы управления в области науки, технологий и инноваций», продеклариро-

ванной в «Стратегии научно-технологического развития Российской Федерации», необходимо разрабатывать новые инструменты мониторинга результативности научной деятельности.



Существующая схема организации мониторинга состояния и результативности научной сферы на основе ежегодного сбора статистической отчётности

Заключение

Научная сфера представляет собой сложную экосистему, включающую научные институты, университеты, R&D-подразделения бизнес-структур, научные фонды, государственные академии наук, большие коллаборации исследователей и отдельных научных лидеров, учёных, изучающих космос или жизнь микроорганизмов, вопросы глобального потепления или фольклор исчезающих народов. Для управления этой сложной системой необходимо глубокое понимание происходящих в ней процессов, основанное на достоверной информации, получаемой при помощи надёжных инструментов. Такую достоверность можно обеспечить с помощью наукометрических методов, основанных на

первичных данных и позволяющих верифицировать процедуру анализа от этапа их получения до конечного результата.

В России, как и в ряде других стран [23], активное применение наукометрии началось с изменения процедур оценки научной результативности и введения наукометрических показателей. Это оказало большое влияние на переосмысление исследователями результатов своего труда и на развитие наукометрии [24]. С одной стороны, общая эффективность российских исследований повысилась (в терминах количества публикаций в журналах, индексируемых в Web of Science и Scopus), а их результаты стали более заметными в мире. С другой стороны, некорректное применение наукометрических методов администраторами науки нередко приводит к некомпетентным решениям, что подрывает доверие к этой научной дисциплине.

Однако наукометрический подход следует рассматривать гораздо шире – как основной инструмент цифровизации научной деятельности и научных коммуникаций. Он предполагает формирование, обработку и анализ массивов данных, содержащих относительно полный состав сведений об исследуемом объекте научной деятельности. Преимущество такого подхода заключается в более высокой достоверности (исследователь контролирует все этапы анализа, начиная с получения первичных данных), воспроизводимости (исследование всегда можно повторить и проверить с использованием тех же данных и методик) и полноте охвата (которая лимитируется лишь ограничениями источников данных). Его результаты могут применяться не только для оценки, но и для более эффективного распространения знания, анализа научного ландшафта, поиска оптимальных решений исследовательских задач.

Весомый вклад в проблематику взаимного развития российской науки и российского книгоиздания с позиций наукометрического (библиометрического) анализа вносит недавно опубликованная, глубокая и детализированная статья В. А. Цветковой [25].

Основной целью научной статьи были формулирование и обоснование вызовов, которые препятствуют дальнейшему развитию и применению наукометрических методов. Преодоление этих вызовов является коллективной задачей научного и издательского сообщества. Для её решения необходимо улучшать практики научных коммуникаций и формировать более чёткий цифровой след исследовательской деятельности.

Список источников

1. **Захарчук Т. В., Грузова А. А.** Научная коммуникация в библиотечно-информационной сфере // Научные и технические библиотеки. 2021. № 3. С. 71–94. doi: 10.33186/1027-3689-2021-3-71-94
2. **Кононова Е. В., Сукиасян Э. Р.** Публикации классификационных систем в интернете: особенности представления и использования // Научные и технические библиотеки. 2021. № 2. С. 91–100. doi: 10.33186/1027-3689-2021-2-91-100
3. **Волкова Н. А., Ходанович М. А.** Классификация гуманитарного знания в ББК: состояние и перспективы // Научные и технические библиотеки. 2021. № 3. С. 31–42. doi: 10.33186/1027-3689-2021-3-31-42
4. **LC Bibliographic Framework Initiative Update Forum** : presentation at Amer. Libr. Assoc. Midwinter conf. (Seattle, Jan. 27, 2013). 2013. URL: <https://www.loc.gov/item/webcast-5789/> (дата обращения: 12.08.2021 г.).
5. **Van de Sompel H., Beit-Arie O.** Open Linking in the Scholarly Information Environment Using the OpenURL Framework // D-Lib Magazine. 2001. Vol. 7. № 3.
6. **Лутай А. В., Любушко Е. Э.** Сравнение качества метаданных в БД CrossRef, Lens, Open-Alex, Scopus, Semantic Scholar, Web of Science Core Collection. URL: https://podpiska.rfbr.ru/storage/reports2021/2022_meta_quality.html (дата обращения: 18.01.2023).
7. **Visser M., Van Eck N. J., Waltman L.** Large-scale comparison of bibliographic data sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic // Quantitative Science Studies. 2021. Vol. 2. № 1. P. 20–41.
8. **Лапочкина В. В., Долгова В. Н., Дикусар К. С., Богатов В. В.** Методический подход к оценке цитируемости научных статей российских журналов в разрезе областей науки по данным Web of Science Core Collection. Часть 1 // Научные и технические библиотеки. 2021. № 4. С. 53–72. doi: 10.33186/1027-3689-2021-4-53-72
9. **Лапочкина В. В., Долгова В. Н., Дикусар К. С., Богатов В. В.** Методический подход к оценке цитируемости научных статей российских журналов в разрезе областей науки по данным Web of Science Core Collection. Часть 2 // Научные и технические библиотеки. 2021. № 5. С. 73–89. doi: 10.33186/1027-3689-2021-5-73-98
10. **Mazov N. A., Gureev V. N.** The role of unique identifiers in bibliographic information systems // Sci. Tech. Inf. Proc. 2014. Vol. 41. № 3. P. 206–210.
11. **Haak L. L. et al.** ORCID: a system to uniquely identify researchers // Learn. Pub. 2012. Vol. 25. № 4. P. 259–264.
12. **Селиванова И. В. и др.** Влияние ошибок в базе данных Scopus на оценку результативности научных исследований // Научно-техническая информация. Серия 1: Организация и методика информационной работы. 2019. № 9. P. 25–32.

13. **Соколов М. М., Титаев К. Д.** Провинциальная и туземная наука // Антропологический форум. 2013. № 19. С. 239–275.
14. **Еременко Т. В.** Оценка международной научной активности в сфере библиотекведческих исследований (2011–2020 гг.) на основе онлайн-платформы SciVal // Научные и технические библиотеки. 2022. № 1. С. 35–56. doi: 10.33186/1027-3689-2022-1-35-56.
15. **Van Eck N. J., Waltman L.** Software survey: VOSviewer, a computer program for bibliometric mapping // *Scientometrics*. 2010. Vol. 84. № 2. P. 523–538.
16. **Chen C.** Science Mapping: A Systematic Review of the Literature // *Journal of Data and Information Science*. 2017. Vol. 2. № 2. P. 1–40.
17. **Стерлигов И. А.** Российский конференционный взрыв: масштабы, причины, дальнейшие действия. Управление наукой: теория и практика. URL: <https://cyberleninka.ru/article/n/rossiyskiy-konferentsionny-vzryv-masshtaby-prichiny-dalneyshie-deystviya> (дата обращения: 18.01.2023).
18. **Vanecsek J., Pecha O.** Fast growth of the number of proceedings papers in atypical fields in the Czech Republic is a likely consequence of the national performance-based research funding system: 3 // *Research Evaluation*. 2020. Vol. 29. № 3. P. 245–262.
19. **Purnell P. J.** Conference proceedings publications in bibliographic databases: a case study of countries in Southeast Asia // *Scientometrics*. 2021. Vol. 126. № 1. P. 355–387.
20. **Bladek M.** DORA: San Francisco Declaration on Research Assessment (May 2013) // *College and Research Libraries News*. 2014. Vol. 75. № 4. P. 191–196.
21. **Hicks D. et al.** Leiden manifesto for research Metrics // Leiden manifesto for research Metrics. URL: <http://www.leidenmanifesto.org/> (дата обращения: 29.06.2021).
22. **Гуськов А. Е., Косяков Д. В.** Проблемы мониторинга научных кадров // *Труды ГПНТБ СО РАН*. 2019. № 1. С. 55–61.
23. **Hicks D.** Performance-based university research funding systems // *Research Policy*. 2012. Vol. 41. № 2. P. 251–261.
24. **Guskov A., Kosyakov D., Selivanova I.** Scientometric research in Russia: impact of science policy changes // *Scientometrics*. 2016. Vol. 107. № 1. P. 287–303.
25. **Цветкова В. А., Мохначева Ю. В.** Российская наука и российское книгоиздание в цифрах и библиометрических оценках // *Научные и технические библиотеки*. 2022. № 11. С. 29–56.

References

1. **Zaharchuk T. V., Gruzova A. A.** Nauchnaia kommunikatsiia v bibliotechno-informatcionnoi` sfere // Nauchny`e i tekhnicheskie biblioteki. 2021. № 3. S. 71–94. doi: 10.33186/1027-3689-2021-3-71-94
2. **Kononova E. V., Sukiasian E` R.** Publikatsii klassifikatsionny`kh sistem v internete: osobennosti predstavleniia i ispol`zovaniia // Nauchny`e i tekhnicheskie biblioteki. 2021. № 2. S. 91–100. doi: 10.33186/1027-3689-2021-2-91-100
3. **Volkova N. A., Hodanovich M. A.** Klassifikatsiia gumanitarnogo znaniia v BBK: sostoiianie i perspektivy` // Nauchny`e i tekhnicheskie biblioteki. 2021. № 3. S. 31–42. doi: 10.33186/1027-3689-2021-3-31-42
4. **LC Bibliographic Framework Initiative Update Forum** : presentation at Amer. Libr. Assoc. Midwinter conf. (Seattle, Jan. 27, 2013). 2013. URL: <https://www.loc.gov/item/webcast-5789/> (data obrashcheniia: 12.08.2021 r.).
5. **Van de Sompel H., Beit-Arie O.** Open Linking in the Scholarly Information Environment Using the OpenURL Framework // D-Lib Magazine. 2001. Vol. 7. № 3.
6. **Lutai` A. V., Liubushko E. E`.** Sravnenie kachestva metadanny`kh v BD CrossRef, Lens, OpenAlex, Scopus, Semantic Scholar, Web of Science Core Collection. URL: https://podpiska.rfbr.ru/storage/reports2021/2022_meta_quality.html (data obrashcheniia: 18.01.2023).
7. **Visser M., Van Eck N. J., Waltman L.** Large-scale comparison of bibliographic data sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic // Quantitative Science Studies. 2021. Vol. 2. № 1. P. 20–41.
8. **Lapochkina V. V., Dolgova V. N., Dikusar K. S., Bogatov V. V.** Metodicheskii` podhod k ocenke tctiruemosti nauchny`kh statei` rossii`skikh zhurnalov v razreze oblastei` nauki po dannym Web of Science Core Collection. Chast` 1 // Nauchny`e i tekhnicheskie biblioteki. 2021. № 4. S. 53–72. doi: 10.33186/1027-3689-2021-4-53-72
9. **Lapochkina V. V., Dolgova V. N., Dikusar K. S., Bogatov V. V.** Metodicheskii` podhod k ocenke tctiruemosti nauchny`kh statei` rossii`skikh zhurnalov v razreze oblastei` nauki po dannym Web of Science Core Collection. Chast` 2 // Nauchny`e i tekhnicheskie biblioteki. 2021. № 5. S. 73–89. doi: 10.33186/1027-3689-2021-5-73-98
10. **Mazov N. A., Gureev V. N.** The role of unique identifiers in bibliographic information systems // Sci. Tech. Inf. Proc. 2014. Vol. 41. № 3. P. 206–210.
11. **Haak L. L. et al.** ORCID: a system to uniquely identify researchers // Learn. Pub. 2012. Vol. 25. № 4. P. 259–264.
12. **Selivanova I. V. i dr.** Vliianie oshibok v baze danny`kh Scopus na ocenku rezul`tativnosti nauchny`kh issledovaniy` // Nauchno-tekhnicheskaiia informatsiia. Serii 1: Organizatsiia i metodika informatcionnoi` raboty`. 2019. № 9. P. 25–32.
13. **Sokolov M. M., Titaev K. D.** Provintcial`naia i tuzemnaia nauka // Antropologicheskii` forum. 2013. № 19. C. 239–275.

14. **Eremenko T. V.** Ocenka mezhdunarodnoi` nauchnoi` aktivnosti v sfere bibliotekovedcheskikh issledovaniï` (2011–2020 gg.) na osnove onlai`n-platformy` SciVal // Nauchny`e i tekhnicheskie biblioteki. 2022. № 1. S. 35–56. doi: 10.33186/1027-3689-2022-1-35-56.
15. **Van Eck N. J., Waltman L.** Software survey: VOSviewer, a computer program for bibliometric mapping // Scientometrics. 2010. Vol. 84. № 2. P. 523–538.
16. **Chen C.** Science Mapping: A Systematic Review of the Literature // Journal of Data and Information Science. 2017. Vol. 2. № 2. P. 1–40.
17. **Sterligov I. A.** Rossiï`skii` konferentsionny`i` vzryv: masshtaby`, prichiny`, dal`nei`shie dei`stviiia. Upravlenie naukoï`: teoriia i praktika. URL: <https://cyberleninka.ru/article/n/rossiyskiy-konferentsionnyy-vzryv-masshtaby-prichiny-dalneyshie-deystviya> (data obrashcheniia: 18.01.2023).
18. **Vanecek J., Pecha O.** Fast growth of the number of proceedings papers in atypical fields in the Czech Republic is a likely consequence of the national performance-based research funding system: 3 // Research Evaluation. 2020. Vol. 29. № 3. P. 245–262.
19. **Purnell P. J.** Conference proceedings publications in bibliographic databases: a case study of countries in Southeast Asia // Scientometrics. 2021. Vol. 126. № 1. P. 355–387.
20. **Bladek M.** DORA: San Francisco Declaration on Research Assessment (May 2013) // College and Research Libraries News. 2014. Vol. 75. № 4. P. 191–196.
21. **Hicks D. et al.** Leiden manifesto for research Metrics // Leiden manifesto for research Metrics. URL: <http://www.leidenmanifesto.org/> (data obrashcheniia: 29.06.2021).
22. **Gus`kov A. E., Kosiakov D. V.** Problemy` monitoringa nauchny`kh kadrov // Trudy` GPNTB SO RAN. 2019. № 1. S. 55–61.
23. **Hicks D.** Performance-based university research funding systems // Research Policy. 2012. Vol. 41. № 2. P. 251–261.
24. **Guskov A., Kosyakov D., Selivanova I.** Scientometric research in Russia: impact of science policy changes // Scientometrics. 2016. Vol. 107. № 1. P. 287–303.
25. **Tcvetkova V. A., Mokhnacheva Iu. V.** Rossiï`skaia nauka i rossiï`skoe knigoizdanie v tcifrah i bibliometricheskikh ocenках // Nauchny`e i tekhnicheskie biblioteki. 2022. № 11. S. 29–56.



Информация об авторах / Information about the authors

Гуськов Андрей Евгеньевич – канд. техн. наук, заведующий лабораторией наукометрии и научных коммуникаций Российского научно-исследовательского института экономики, политики и права в научно-технической сфере, Москва, Российская Федерация
guskov.andrey@gmail.com

Шрайберг Яков Леонидович – доктор техн. наук, профессор, член-корреспондент Российской академии образования, научный руководитель ГПНТБ России, главный редактор журнала «Научные и технические библиотеки», Москва, Российская Федерация; заведующий кафедрой электронных библиотек и наукометрических исследований Московского государственного лингвистического университета, Москва, Российская Федерация
gpntb@gpntb.ru

Andrey E. Guskov – Cand. Sc. (Engineering), Head, Laboratory for Scientometrics and Scholarly Communications, Russian Research Institute of Economics, Politics and Law in Science and Technology, Moscow, Russian Federation
guskov.andrey@gmail.com

Yakov L. Shrayberg – Dr. Sc. (Engineering), Professor, Corresponding Member of Russian Academy of Education; Director of Research, Russian National Public Library for Science and Technology, Editor-in-Chief, “Scientific and Technical Libraries” Journal, Moscow, Russian Federation; Head, Department for Electronic Libraries and Scientometric Studies, Moscow State Linguistic University, Moscow, Russian Federation
gpntb@gpntb.ru