

УДК 01, 002.53

РЕЙТИНГ ЖУРНАЛА В БИБЛИОГРАФИЧЕСКОЙ БАЗЕ

М. М. Горбунов-Посадов¹ [0000-0002-7044-8287], Т. А. Полилова² [0000-0003-4628-3205]

^{1, 2}Институт прикладной математики им. М.В. Келдыша Российской академии наук, Миусская пл., 4, Москва, 125047

¹gorbunov@keldysh.ru, ²polilova@keldysh.ru

Аннотация

Инструмент построения рейтингов научных журналов является одним из востребованных сервисов библиографических баз. Задача построения рейтинга обычно делится на две основные подзадачи: определение референтной группы журналов и вычисление показателя рейтинга для журналов этой группы. Практика показывает, что для корректного сопоставления журналов необходимым условием является ограничение референтной группы исключительно журналами определенной тематики. В случае методических ошибок, допущенных на этапе выделения референтной группы, значения показателя журналов в рейтинге могут сильно отличаться от ожидаемых.

Например, в рейтинге журналов в Российском индексе научного цитирования (РИНЦ) по двухлетнему импакт-фактору в тематическом направлении «Математика» классические фундаментальные математические журналы вопреки ожиданиям не выходят на первые позиции рейтинга. Первые позиции заняли журналы, для которых математика не является доминирующей профильной дисциплиной. Анализ статистических данных о тематике публикуемых статей и цитирований в журналах, занимающих лидирующие позиции рейтинга РИНЦ, показывает, что на показатели рейтинга существенно повлияла мультидисциплинарность этих журналов.

Отмеченное недоразумение подводит к мысли о том, что в подсчет рейтинга в данном случае следовало вовлекать не все статьи журнала, а только относящиеся к данному тематическому направлению. Вместе с тем вопросы вызывает и сложившаяся схема тематической классификации направлений. Более перспективной представляется набирающая популярность классификация «снизу вверх», работающая на представительном массиве статей. Здесь тематиче-

ские кластеры вычленяются на основе понятия близости статей, трактуемого как близость их библиографических связей. И далее тематическая принадлежность статьи не назначается волевым решением автора или редакции, а строго формально вычисляется на основе ее библиографического списка.

***Ключевые слова:** научная публикация, цитирование, рейтинг журналов, тематическая классификация, импакт-фактор, мультидисциплинарность, библиографическая ссылка, со-цитирование, классификация снизу вверх, тематическая кластеризация, Citation Topics.*

ВВЕДЕНИЕ

Современные библиографические базы накопили огромный массив информации о научных публикациях. Научные издательства регулярно поставляют в библиографические базы широкий набор метаданных издаваемой научной продукции. В библиографические базы попадают данные о журналах, названия и аннотации статей, ключевые слова, коды тематических направлений, данные об авторах статей и организациях, аффилированных с авторами. В базы загружаются библиографические списки литературы, используемой в статье, которые преобразуются в формальные структуры. Эти структуры допускают машинную обработку и обеспечивают отождествление элемента библиографического списка со статьей. В последние годы благодаря международным кодам DOI, ORCID, ROR качество обработки библиографических списков и процедур отождествления библиографических ссылок со статьями заметно улучшилось.

Развитые библиографические базы предоставляют своим многочисленным пользователям интернет-доступ к различным инструментам для поиска журналов, статей и авторов, интересующих читателя. На странице автора, найденной через поисковый сервис базы, любой посетитель может увидеть все публикации данного автора, размещенные в этой базе. Автору становится доступной информация о числе цитирований каждой его статьи. Доступны также ссылки на статьи, цитирующие публикации автора. Пользователь может осуществить поиск интересующего журнала или группы журналов по указанным поисковым атрибутам. Перейдя на страницу журнала, посетитель получает более подробные сведения о журнале и осуществляемой им редакционной политике.

Карточка журнала содержит информацию о тематике журнала — обычно приводятся тематические коды из используемого рубрикатора.

Данные, хранящиеся в библиографической базе, предоставляют пользователю статистические сведения и наукометрические показатели журнала, в частности: число опубликованных статей и цитирований статей журнала, позиции журнала в тематических рейтингах, распределение публикаций по авторам или организациям, по тематике публикуемых или цитирующих статей, распределение цитирующих публикаций по журналам или организациям и т. д.

Библиографическая база становится мощным инструментом, позволяющим пользователю, не имеющему специальной подготовки в области наукометрии, проводить несложные аналитические исследования. Как отмечено в [1], появилась новая целевая аудитория, которая не занимается наукометрическими исследованиями на профессиональном уровне, но проявляет большой интерес к наукометрическим показателям и рейтингам научных журналов. К этой аудитории относятся сотрудники организаций и фондов, финансирующих научные исследования, а также научные работники. Обладая общей информационной культурой, зная методологические принципы проведения научных исследований, имея опыт анализа и обработки данных, специалисты из разных научных областей активно интерпретируют результаты наукометрических исследований, получаемые профессионалами.

Следует отметить, что и в среде профессиональных специалистов в области наукометрии до сих пор не утихают споры по фундаментальным методологическим вопросам: идет поиск новых индикаторов, способствующих более адекватному сравнению показателей журналов, новых подходов, обеспечивающих аккуратную нормализацию цитируемости по разным предметным областям на уровне журналов и на уровне отдельных статей [2]. Приходит понимание, что наукометрические исследования должны уделять больше внимания географическим, социальным и языковым измерениям, что приведет к расширению набора показателей и аналитических приемов, используемых для оценки эффективности исследований [3].

Одним из наиболее востребованных сервисов библиографических баз является инструмент построения рейтингов журналов по выбранным показателям,

например, по известному широкой научной аудитории показателю импакт-фактора. Ученые, являясь заинтересованными потребителями результатов ранжирования журналов, в состоянии оценить адекватность рейтингов в своей научной области, построенных профессиональными наукометристами. Ученые имеют личный опыт подготовки и издания научных статей в журналах по своей тематике и, безусловно, знают не понаслышке авторитетные журналы своего научного направления.

Задача построения тематического рейтинга обычно делится на два этапа:

- выбор референтной группы журналов в исследуемом тематическом направлении,
- вычисление показателей импакт-фактора для журналов референтной группы на выбранном поле цитирующих журналов в определенном временном интервале.

Результаты рейтинга зависят от методических приемов и алгоритмических решений, которые принимаются на каждом из перечисленных этапов. Анализ методик построения рейтингов научных журналов в некоторых случаях позволяет выявить источники перекосов, ошибок или неточностей, обнаружить групповые интересы или предвзятости составителей рейтинга. Поэтому важными основополагающими принципами, сформулированными в Лейденском Манифесте наукометрии [4], объявлены открытость процедур сбора и анализа данных для проведения наукометрических оценок, а также возможность самим ученым-исследователям проверять правильность определения результатов наукометрического анализа.

Известно, что если поведение построенной математической модели не соответствует физическим реалиям, то теоретические положения, ограничения или допущения в математической модели требуют дополнительного тщательного анализа. Аналогично, если результаты тематического рейтинга научных журналов не отвечают ожиданиям специалистов данной научной области, следует проанализировать не только формулу вычисления рейтинга (она может быть весьма простой), но и методику отбора референтной группы, принятые ограничения (расширения) множества цитирующих журналов, способы учета особенностей цитируемости в различных научных областях.

1. РЕЙТИНГИ ЖУРНАЛОВ В РОССИЙСКОМ ИНДЕКСЕ НАУЧНОГО ЦИТИРОВАНИЯ

Рассмотрим инструменты, которые предоставляет Российский индекс научного цитирования (РИНЦ) [5], функционирующий на платформе библиографической базы eLibrary.ru [6]. В настоящее время в РИНЦ индексируется более 5700 научных журналов.

РИНЦ формирует рейтинги журналов по таким показателям как:

- число цитирований,
- двухлетний и пятилетний импакт-факторы с возможными ограничениями по типам научных изданий и области цитирования,
- индекс Херфиндаля по цитирующим журналам или организациям авторов,
- индекс Хирша

и другим наукометрическим показателям.

С помощью экранной формы РИНЦ можно построить общий рейтинг журналов по перечисленным показателям, а также тематические рейтинги для журналов определенного научного направления.

Построим рейтинг РИНЦ «Двухлетний импакт-фактор РИНЦ» по тематическому направлению «Математика» за 2019 год (рис. 1). В этом рейтинге при расчете показателей импакт-фактора учитываются цитирования из всех журналов, включенных в РИНЦ: первые 10 позиций занимают журналы, представленные в Таблице 1.

Результаты рейтинга, представленные в таблице, не вполне соответствуют ожиданиям ученых-математиков. Существующий математический портал Math-Net.ru [7] дает более адекватную картину на поле математических журналов. Наиболее известными и авторитетными в среде математиков являются журналы математической тематики, издаваемые Российской академией наук (РАН), Математическим институтом им. В.А. Стеклова РАН, Российской академией наук, Отделением математических наук и другими авторитетными академическими структурами.

В приведенном выше рейтинге РИНЦ обращает на себя внимание лидирующая позиция журнала «Геометрия и графика», который не относится к топовым

математическим журналам. Чтобы понять, в чем причина столь высоких показателей этого журнала в направлении «Математика», рассмотрим более внимательно анкету журнала и статистические отчеты о тематике издаваемых статей и тематике цитирующих публикаций.

Таблица 1

**Рейтинг по двухлетнему импакт-фактору РИНЦ
для журналов в тематическом направлении «Математика» за 2019 год
(позиции рейтинга 1–10)**

№ в рейтинге ИФ-2 РИНЦ	Название журнала	Значение показателя в рейтинге 2-ИФ РИНЦ
1.	Геометрия и графика	1,899
2.	Информатика и автоматизация	1,684
3.	Вестник Самарского государственного технического университета. Серия: Физико-математические науки	1,024
4.	Известия Российской академии наук. Серия математическая	0,978
5.	Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления	0,957
6.	Известия Иркутского государственного университета. Серия: Математика	0,932
7.	Вычислительная механика сплошных сред	0,760
8.	Математические заметки	0,757
9.	Экономика и математические методы	0,750
10.	Pattern Recognition and Image Analysis (Advances in Mathematical Theory and Applications)	0,744

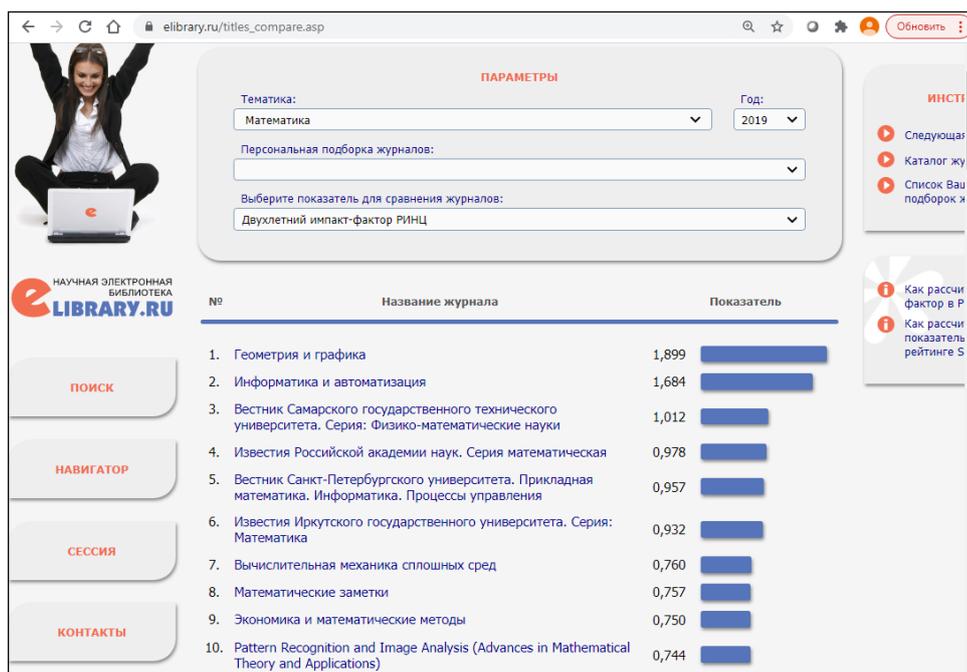


Рис. 1. Рейтинг по показателю «Двухлетний импакт-фактор РИНЦ» по тематическому направлению «Математика» за 2019 год.

Журнал «Геометрия и графика»

Учредителем журнала является частное лицо — В.И. Вышнепольский.

В анкете журнала указано, что этот научный журнал посвящен проблемам геометрии, черчения, компьютерной графики, преподаванию графических дисциплин и других тем, связанных с геометрией и графикой. Важным направлением является исследование отраслевых особенностей применения геометрии и компьютерной графики в строительстве, машиностроении, разработке программного обеспечения и т. д.

Журнал указал следующие тематические рубрики.

РУБРИКИ ГРНТИ:

143500. Высшее профессиональное образование. Педагогика высшей профессиональной школы

272100. Геометрия

281700. Теория моделирования

РУБРИКИ OECD:

101. Mathematics

503. Educational sciences

СПЕЦИАЛЬНОСТИ ВАК:

050100. Инженерная геометрия и компьютерная графика

Рассмотрим распределение статей журнала по тематике (рис. 2).

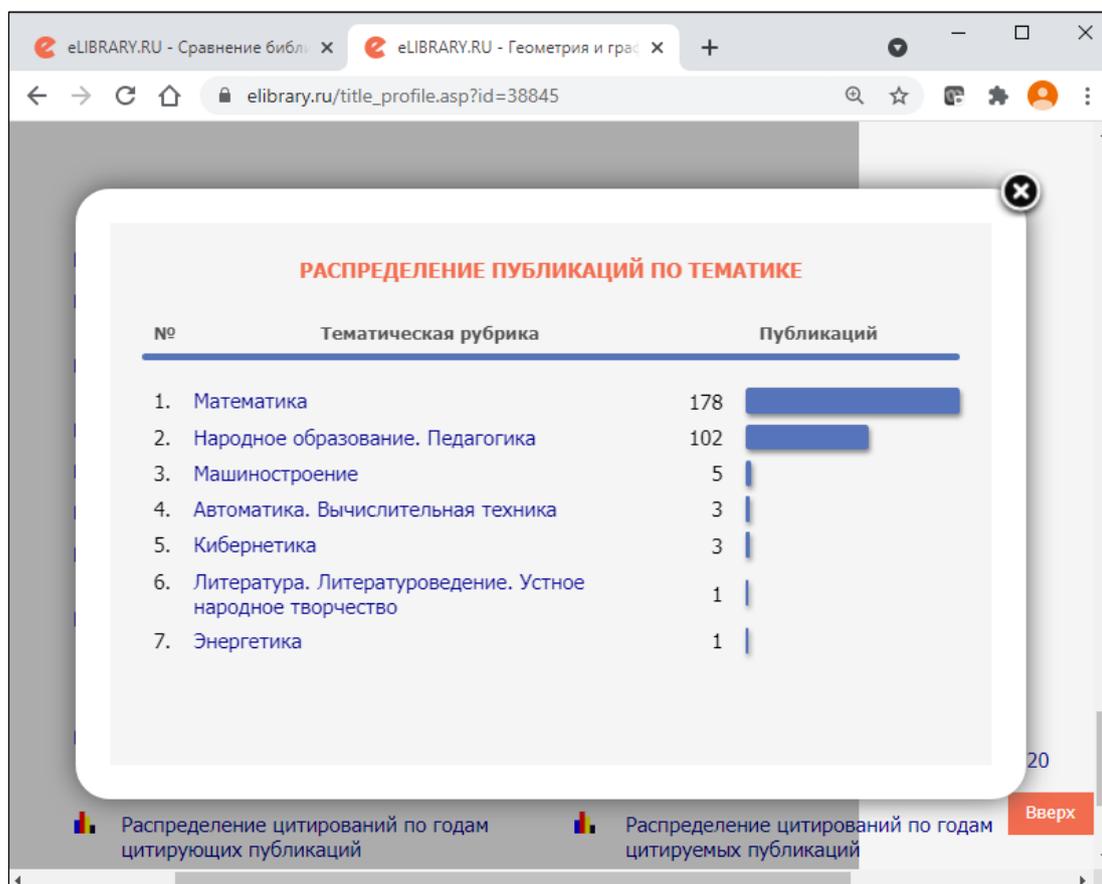


Рис. 2. Статистический отчет РИНЦ «Распределение публикаций по тематике» журнала «Геометрия и графика».

Будем считать, что в этом статистическом отчете, как и в аналогичных последующих, каждая статья заявлена по одной (главной) тематике. В то же время статья, вообще говоря, может быть приписана к нескольким тематическим направлениям. Как в этом случае будет выглядеть отчет? Поскольку разработчики не дают точного описания содержания отчета, предположение о множественном присутствии одной статьи в отчете приходится не принимать во внимание. Однако этой неопределенностью можно пренебречь, поскольку нас в первую очередь будут интересовать не столько точные подсчеты, сколько качественные оценки.

Всего статей, опубликованных в журнале «Геометрия и графика» — 293. По направлению «Математика» в отчете присутствуют 178 статей, т. е. 61% всех статей. Однако, если обратиться к распределению цитирующих публикаций по тематике для журнала «Геометрия и графика» (рис. 3), обнаружатся иные пропорции.

Всего в отчете зафиксировано 832 цитирующие статьи. Цитирующих статей по теме «Математика» — 290 (35%). Как показано в работе [8], высокие показатели цитируемости журналу обеспечили статьи, относящиеся к тематическому направлению «Народное образование. Педагогика» (около 49% всех цитирующих статей). В РИНЦ предусмотрен прицельный отчет «Распределение цитирований по тематике цитирующих публикаций», который дал бы более полную картину по тематике цитирующих публикаций, но, к сожалению, этот отчет оказался недоступным.

Журнал «Геометрия и графика» в рейтинге по двухлетнему импакт-фактору РИНЦ 2019 года в направлении «Народное образование. Педагогика» занимает 10-е место. В этом направлении показатели цитируемости журналов значительно выше, чем в направлении «Математика». Так, лидер рейтинга журнал «Вестник Мининского университета» (Нижний Новгород) имеет показатель двухлетнего импакт-фактора 5,336. Показатели ведущих математических журналов значительно скромнее. Таким образом, высокий показатель журнала «Геометрия и графика» в рейтинге по импакт-фактору в направлении «Математика», как и следовало ожидать, во многом обеспечили цитирования из статей, не относящихся к математике.

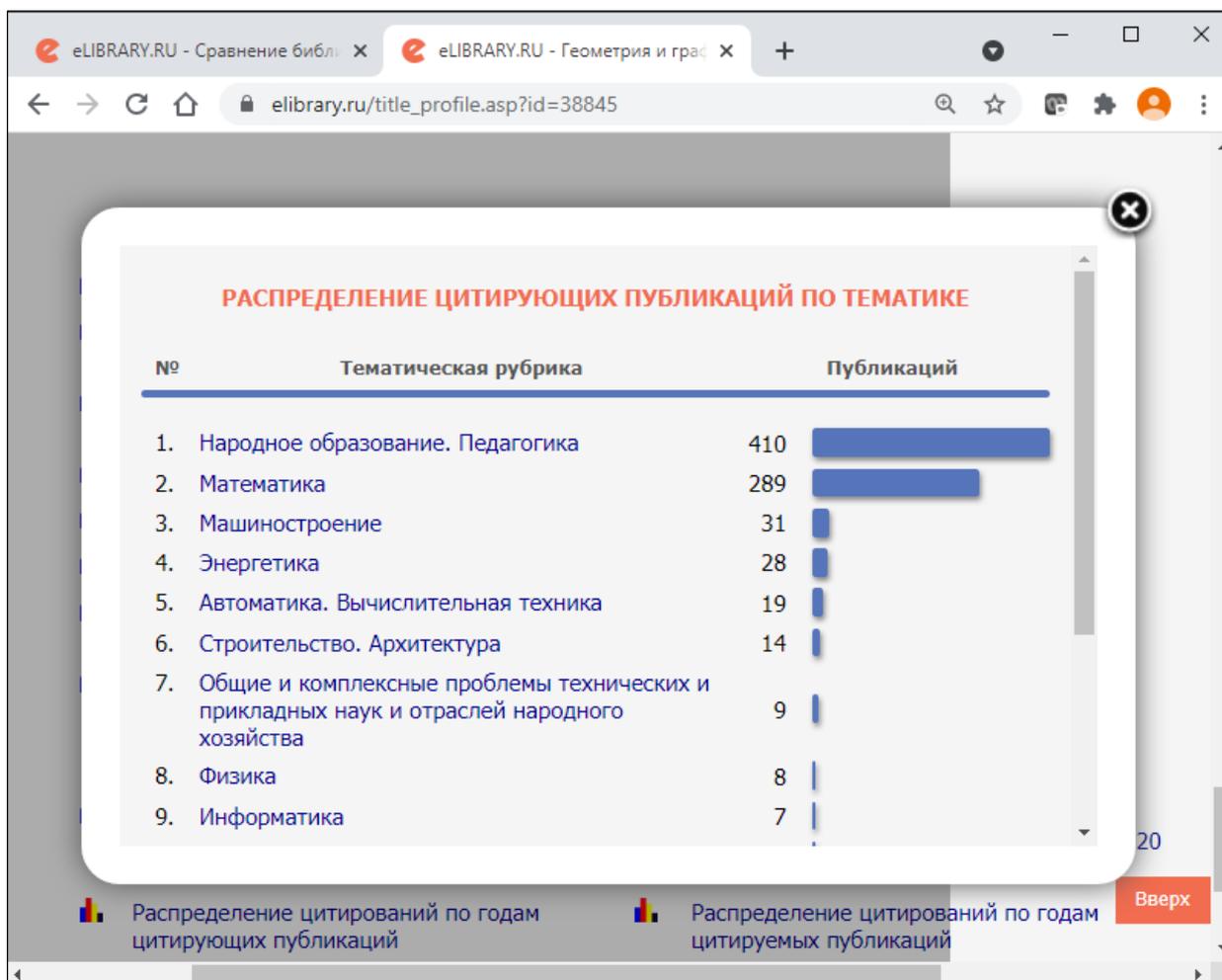


Рис. 3. Статистический отчет РИНЦ «Распределение цитирующих публикаций по тематике» журнала «Геометрия и графика».

В описании методики построения рейтингов РИНЦ по импакт-фактору нет каких-либо сведений об учете специфики цитирования в разных тематических направлениях. Также не сообщается, что методика предусматривает соответствующее нормирование показателей цитирования на уровне статей.

В результате рассмотрения статистических данных о тематике цитирующих статей можно сделать заключение, что на импакт-фактор журнала «Геометрия и графика» существенно повлияла мультидисциплинарность журнала. Журнал достаточно точно определил свои тематические рубрики, например, по классификатору OECD (“Mathematics”, “Educational sciences”). Однако на незаслуженно высокую позицию в рейтинге по импакт-фактору в направлении «Математика»

журнал выводят не относящиеся к делу ссылки из статей по тематике «Народное образование. Педагогика».

Возможен и другой подход к построению рейтинга. В работе [9] выдвинут тезис о том, что оценку влияния *автора* публикаций в некоторой научной области следует проводить на основе показателей цитируемости публикаций, относящихся к этой научной области. Например, авторитет автора в научном направлении «Физика» определяется на основе подсчета цитирований его публикаций по физике, и при этом не следует включать в расчет цитирования публикаций этого автора по другим научным направлениям. Эту идею можно распространить и на оценку влияния *журнала* в разных тематических направлениях. Если бы в подсчете рейтинга в разделе «Математика» учитывались массив статей только по теме «Математика» и цитирования только этих статей, то показатели у журнала «Геометрия и графика», по-видимому, были бы скромнее. И такая методика подсчета точнее определяла бы место журнала «Геометрия и графика» в рейтинге по математическому направлению.

Журнал «Информатика и автоматизация»

Издателем журнала является Санкт-Петербургский Федеральный исследовательский центр РАН. Анкета журнала декларирует, что он является научным, научно-образовательным, междисциплинарным журналом с базовой специализацией в области информатики, автоматизации и прикладной математики.

Журнал отнесен к следующим тематическим направлениям.

РУБРИКИ ГРНТИ:

500000. Автоматика. Вычислительная техника

270000. Математика

280000. Кибернетика

282300. Искусственный интеллект

РУБРИКИ OECD:

101. Mathematics

102. Computer and information sciences

202. Electrical engineering, electronic engineering

СПЕЦИАЛЬНОСТИ ВАК:

- 010102. Дифференциальные уравнения, динамические системы и оптимальное управление
- 010105. Теория вероятностей и математическая статистика
- 010109. Дискретная математика и математическая кибернетика
- 051301. Системный анализ, управление и обработка информации (по отраслям)
- 051311. Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей
- 051315. Вычислительные машины, комплексы и компьютерные сети
- 051317. Теоретические основы информатики
- 051319. Методы и системы защиты информации, информационная безопасность

Журнал занимает 1-е место в интегральном рейтинге SCIENCE INDEX за 2020 год по тематике «Автоматика. Вычислительная техника» и 1-е место по тематике «Кибернетика». В интегральном рейтинге SCIENCE INDEX по направлению «Математика» журнал занимает 2-е место.

Рассмотрим статистический отчет журнала по тематике публикуемых статей (рис. 3). Данные отчета показывают, что доминирующим тематическим направлением журнала является направление Computer Science, включающее тематики «Автоматика. Вычислительная техника», «Кибернетика», «Информатика». Именно по этим трем тематикам опубликовано больше всего статей — 694. По теме «Математика» опубликовано 270 статей.

На рис. 4 представлен статистический отчет о распределении цитирующих статей по тематике.

Доля статей по теме «Математика» от числа статей по темам Computer Science (трех рубрикам «Автоматика. Вычислительная техника», «Кибернетика» и «Информатика») составляет 39%. Доля цитирующих статей по теме «Математика» по отношению к числу цитирующих статей по темам Computer Science составляет только 22%. Таким образом, в журнале «Информатика и автоматизация» математические статьи цитируются хуже, чем статьи по темам Computer Science. Этот вывод совпадает с общепринятыми представлениями о более скромном цитировании статей по математике. Следовательно, можно утвер-

ждать, что учет статей по темам Computer Science и их цитирований заметно повышает показатель журнала «Информатика и автоматизация» в рейтинге по импакт-фактору в направлении «Математика».

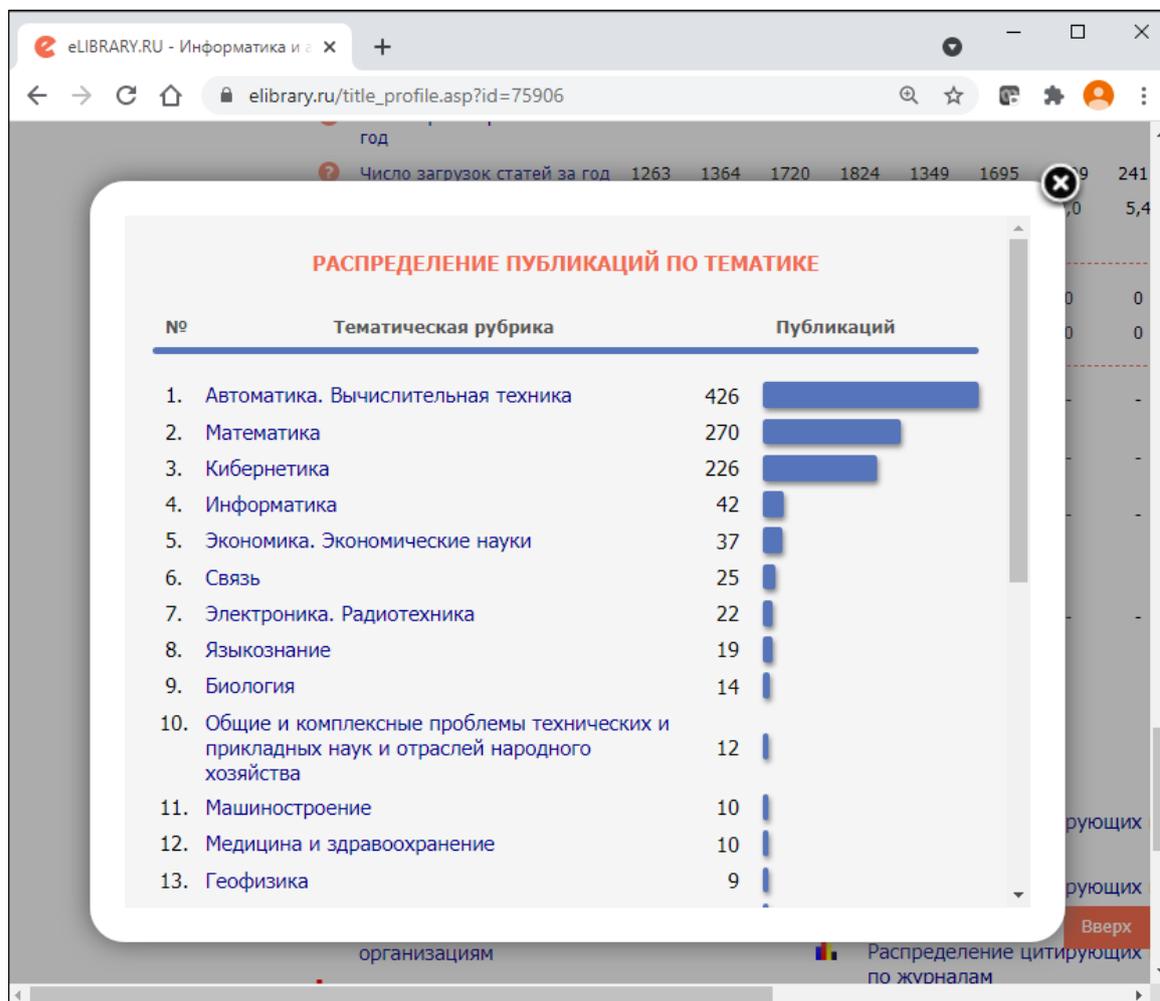


Рис. 3. Статистический отчет РИНЦ «Распределение публикаций по тематике» журнала «Информатика и автоматизация».

Если бы при расчете рейтинга журналов по импакт-фактору в разделе «Математика» учитывались только статьи по теме «Математика» и их цитирования и не учитывались статьи и цитирования по другим темам (в частности, по темам Computer Science), то показатель импакт-фактора по теме «Математика» журнала «Информатика и автоматизация» был бы ниже.

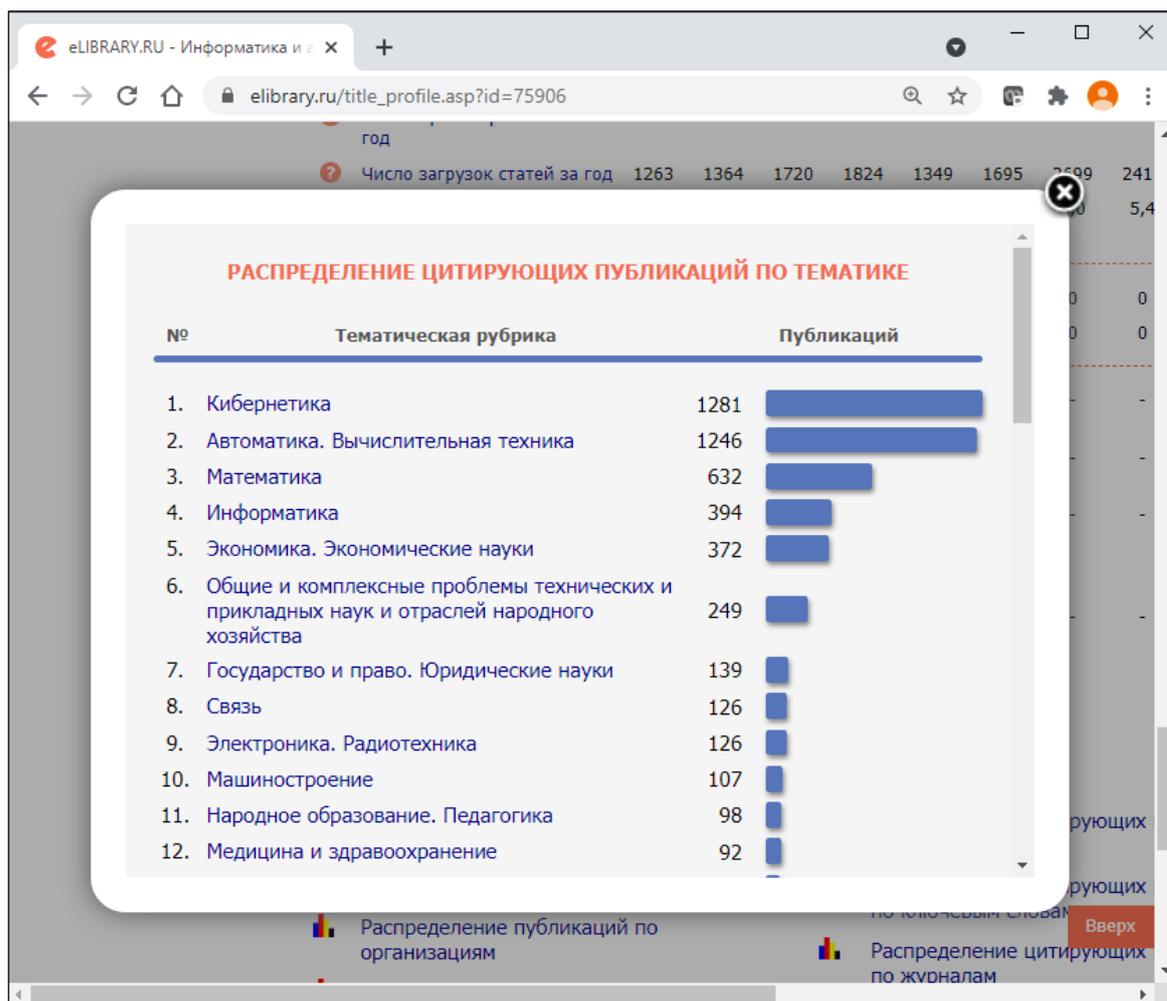


Рис. 4. Статистический отчет РИНЦ «Распределение цитирующих публикаций по тематике» журнала «Информатика и автоматизация».

Журнал «Вестник Самарского государственного технического университета. Серия: Физико-математические науки»

Журнал публикует оригинальные статьи и заказные обзоры по направлениям «Дифференциальные уравнения и математическая физика», «Механика деформируемого твёрдого тела», «Математическое моделирование, численные методы и комплексы программ».

РУБРИКИ ГРНТИ:

272900. Обыкновенные дифференциальные уравнения

273100. Дифференциальные уравнения с частными производными

273300. Интегральные уравнения

273500. Математические модели естественных наук и технических наук.

Уравнения математической физики

274100. Вычислительная математика

301900. Механика деформируемого твердого тела

305100. Комплексные и специальные разделы механики

301700. Механика жидкости и газа

281500. Теория систем автоматического управления

281700. Теория моделирования

РУБРИКИ OECD:

101. Mathematics

203. Mechanical engineering

205. Materials engineering

СПЕЦИАЛЬНОСТИ ВАК:

010102. Дифференциальные уравнения, динамические системы и оптимальное управление

010204. Механика деформируемого твердого тела

051318. Математическое моделирование, численные методы и комплексы программ

В анкете журнала в РИНЦ отмечено, что Вестник не является мультидисциплинарным журналом, но, по-видимому, эта декларация нужна для того, чтобы не относить Вестник к группе мультидисциплинарных журналов РИНЦ, выделенных в отдельную категорию. Формальных критериев отнесения журнала к мультидисциплинарной группе РИНЦ не формулирует. Можно построить в РИНЦ рейтинг для мультидисциплинарных журналов, выделенных в отдельную группу. Обращает на себя внимание тот факт, что в этот рейтинг будут включены журналы, публикующие статьи в широком диапазоне естественно-научных, технических, гуманитарных и общественных направлений. Вестник, вообще говоря, не вписывается в такую группу мультидисциплинарных журналов.

Рассмотрим статистический отчет журнала по тематике публикуемых статей (рис. 5). В журнале доминируют статьи по трем темам: «Математика» «Механика» и «Физика». Больше всего статей в журнале опубликовано по теме «Ма-

тематика» — 639. Число статей по темам «Механика» и «Физика» в сумме не-
много больше, чем число статей по математике — 656.

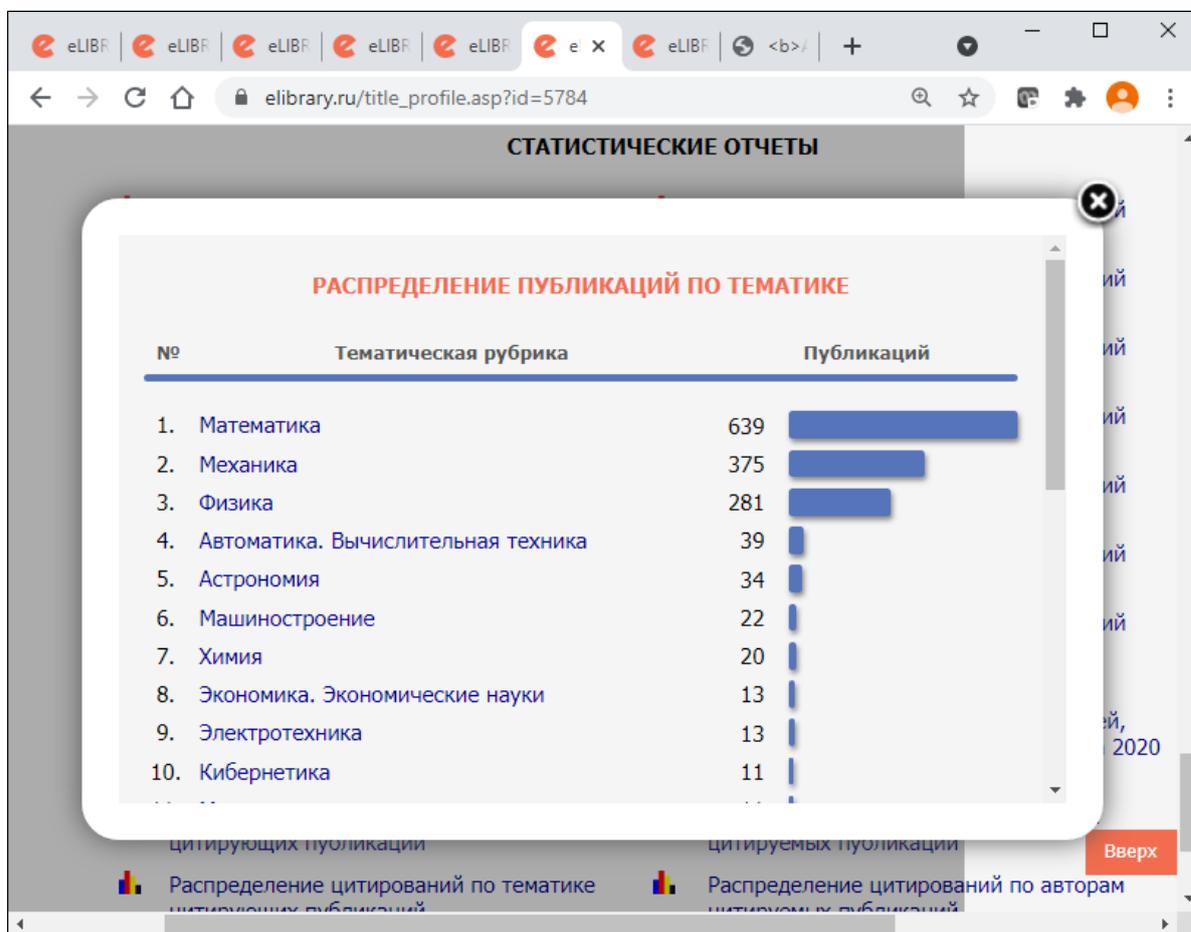


Рис. 5. Статистический отчет РИНЦ «Распределение публикаций по тематике»
журнала «Вестник Самарского государственного технического университета.
Серия: Физико-математические науки».

Рассмотрим статистический отчет по тематике цитирующих статей (рис. 6).

Доминирующими ожидаемо стали следующие три тематических направ-
ления: «Математика», «Механика» и «Физика». Тематическое направление «Фи-
зика» не прозвучало в декларациях редакции на сайте журнала, но этому факту
есть простое объяснение. Математика и физика в научной среде считаются
близкими направлениями. Физики активно используют математический аппарат
для получения фундаментальных и прикладных результатов. Математики разви-
вают математические методы для решения физических задач.

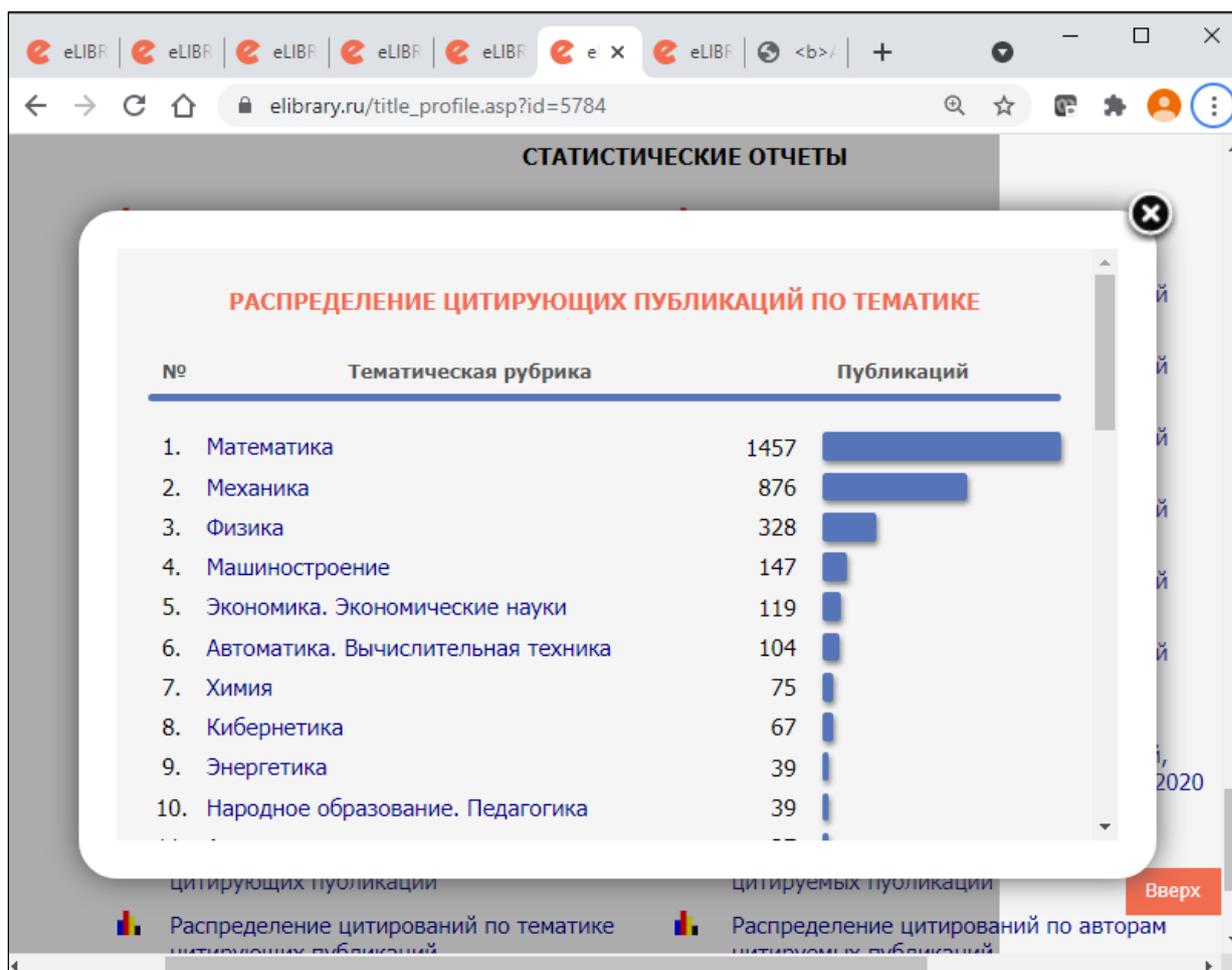


Рис. 6. Статистический отчет РИНЦ «Распределение цитирующих публикаций по тематике» журнала «Вестник Самарского государственного технического университета. Серия: Физико-математические науки».

Если перевести фокус на проблемы тематических рейтингов, то можно отметить более высокую цитируемость статей по направлению «Физика» по сравнению с цитируемостью статей по направлению «Математика». Но, скорее всего, цитируемость статей журнала в направлении «Физика» сильно зависит от конкретной темы исследований.

Появление Вестника на третьей позиции тематического рейтинга в направлении «Математика» не вызывает удивления или резкого отторжения. Однако и здесь более адекватные показатели рейтинга в направлении «Математика» могут быть получены в случае, если при расчете показателя импакт-фактора учитывать только статьи по теме «Математика» и цитирования этих статей.

Журнал «Известия Российской академии наук. Серия математическая»

Издателями журнала являются Математический институт им. В.А. Стеклова РАН (Москва) и Российская академия наук. Журнал публикует статьи по всем разделам современной математики. Особое внимание уделяется алгебре, математической логике, теории чисел, математическому анализу, геометрии, топологии, дифференциальным уравнениям. Редакция заявляет раздел «Математика» по трем классификаторам: ГРНТИ (категория 270000), OECD (категория 101. Mathematics), специальности ВАК (категория 010100.)

На рис. 7 представлен статистический отчет РИНЦ о тематике статей, публикуемых в журнале, на основе собранных данных за весь период его размещения в eLibrary.

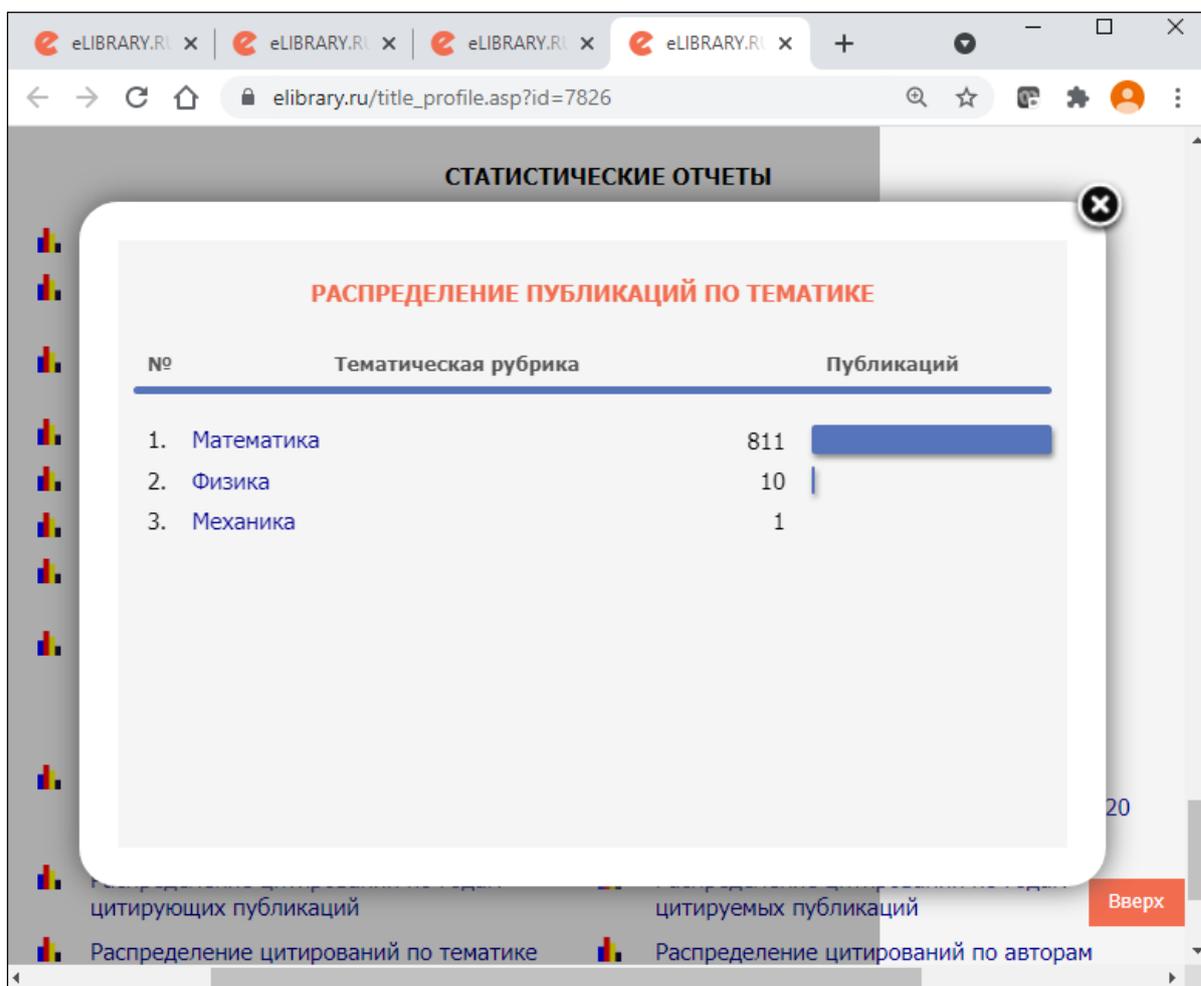


Рис. 7. Статистический отчет РИНЦ «Распределение публикаций по тематике» журнала «Известия Российской академии наук. Серия математическая».

Из представленных статистических данных следует, что практически все статьи, публикуемые в журнале, относятся к теме «Математика»: по этой теме опубликовано 811 статей, по теме «Физика» — 10, и по теме «Механика — 1.

Если посмотреть статистический отчет по цитирующим статьям, то вновь можно убедиться в том, что по теме «Математика» зафиксировано подавляющее большинство статей — 18734. На второй позиции отчета указаны 3483 статьи по теме «Физика», по теме «Автоматика. Вычислительная техника» — 241, теме «Химия» — 210, по теме «Механика» — 182. По остальным тематикам, присутствующим в отчете, зафиксировано меньше сотни цитирующих статей.

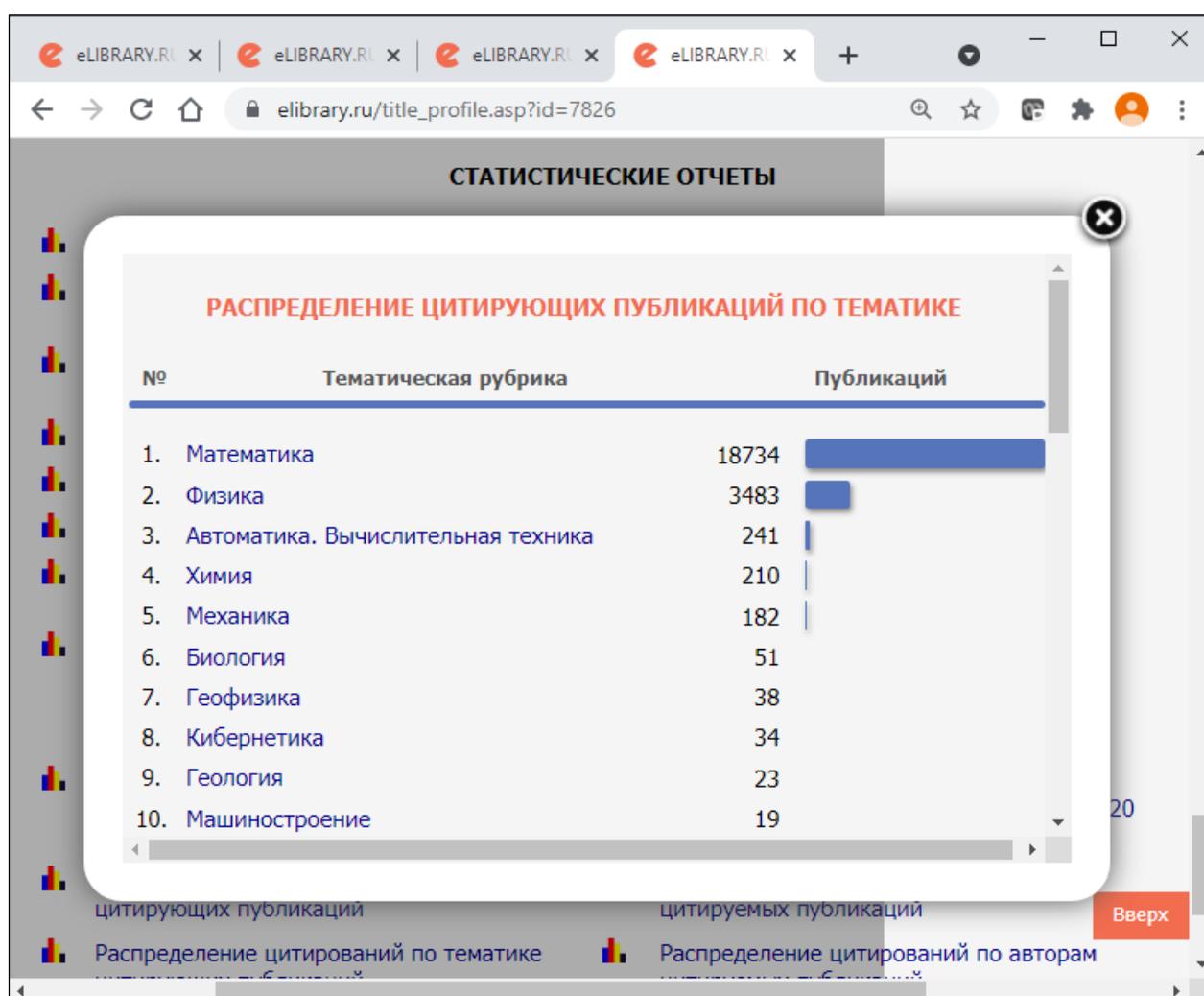


Рис. 8. Статистический отчет РИНЦ «Распределение цитирующих публикаций по тематике» журнала «Известия Российской академии наук. Серия математическая».

Подведем итог рассмотрения статистических данных по тематике публикуемых и цитирующих статей для четырех журналов, занявших лидирующие позиции в рейтинге РИНЦ по двухлетнему импакт-фактору в направлении «Математика».

Первые две позиции в рейтинге по двухлетнему импакт-фактору, занятые журналами «Геометрия и графика» и «Информатика и автоматизация», очевидно, не соответствуют ожиданиям специалистов из области «Математика». Причина столь высоких показателей этих журналов состоит в том, что при расчете показателей рейтинга учитывались статьи, не относящиеся к направлению «Математика», и цитирования этих статей. Рассмотренные два журнала обладают выраженной мультидисциплинарностью в том смысле, что помимо направления «Математика» в журналах дополнительно присутствуют другие, хорошо цитируемые научные дисциплины. В обоих журналах математика не является доминирующей дисциплиной.

Журнал «Вестник Самарского государственного технического университета. Серия: Физико-математические науки» также является мультидисциплинарным: в журнале были опубликованы в большом количестве статьи по двум другим дисциплинам — «Физика», «Механика». Но доминирующим тематическим направлением является «Математика». Эффект влияния на показатели импакт-фактора статей и цитирований, относящихся к двум дисциплинам («Физика», «Механика»), не столь очевиден. Более точно определить влияние этих двух направлений можно только при получении доступа к данным о цитированиях по каждому тематическому направлению.

Журнал «Известия Российской академии наук. Серия математическая» можно назвать классическим математическим журналом: число статей по второй тематике (физике) в этом журнале незначительно. Журнал хорошо известен в кругу профессиональных математиков. Этот журнал вполне мог бы претендовать на лидирующую позицию в тематическом рейтинге по направлению «Математика». Но в рассмотренном рейтинге журналу приходится соревноваться с мультидисциплинарными журналами, включающими статьи по тематикам, которые традиционно имеют более высокие показатели цитируемости.

2. ПРОБЛЕМА ВЫБОРА РЕФЕРЕНТНОЙ ГРУППЫ ЖУРНАЛОВ

На наш взгляд, причина выхода непрофильных журналов на первые позиции в рейтинге по направлению «Математика» кроется в методике формирования референтной группы журналов. Референтная группа — это множество «похожих» журналов, относящихся к общей тематике, определяемой тематическим классификатором базы. Как поступать в том случае, когда журнал содержит статьи по другим тематикам, не относящимся к тематике рейтинга?

Известно, что в РИНЦ ведена категория «Мультидисциплинарные журналы». По-видимому, к этой категории отнесены журналы, которые декларируют мультидисциплинарность в своей редакционной политике и инициативно объявляют себя «мультидисциплинарными». Такие журналы, как правило, ориентированы на широкий круг научных направлений, относящихся к тематическим категориям верхнего уровня классификатора (такую мультидисциплинарность условно можно назвать «сильной»).

Рассмотренные нами журналы имеют разную степень мультидисциплинарности. Журнал «Геометрия и графика» можно отнести к журналам с «сильной» мультидисциплинарностью, поскольку он содержит две доминирующие тематики из разных категорий верхнего уровня классификатора. Журналы «Информатика и автоматизация» и «Вестник Самарского технического университета» можно отнести к журналам, имеющим «среднюю» мультидисциплинарность. Классический математический журнал «Известия Российской академии наук. Серия математическая» относится в предложенной шкале скорее к журналам со «слабо выраженной» мультидисциплинарностью. В рейтинге РИНЦ по двухлетнему импакт-фактору первые три позиции заняли журналы с высокой и средней степенями мультидисциплинарности.

Мультидисциплинарность зависит от выбранной системы тематической классификации. На рассмотренных примерах мы видим, что разработчики РИНЦ опираются на несколько систем классификации:

- ГРНТИ — российский Государственный рубрикатор научно-технической информации, представляющий собой универсальную иерархическую классификацию областей знания, принятую для систематизации научно-технической информации в России и государствах СНГ,

- OECD — международный классификатор OECD (Organization for Economic Cooperation and Development),
- номенклатуру специальностей Высшей аттестационной комиссии (ВАК),

а также на свой собственный рубрикатор, сочетающий в себе подходы перечисленных классификаторов, адаптированные для решения конкретных потребностей информационных сервисов. Вероятно, в РИНЦ немалую роль играют экспертные оценки и особые решения при отнесении журнала к той или иной тематической категории. К сожалению, РИНЦ не предоставляет пользователю достаточных справочных материалов с обоснованием выбора или модификации конкретного классификатора и описанием применяемых методик формирования референтных групп.

Существует проблема, связанная со структурой используемого классификатора. Приведем пример. При расчете показателей журналов в некоторых типах рейтингов РИНЦ в одну рубрику попадают два тематических класса: “Mathematics” и “Computer and information sciences”. В классификаторе OECD эти направления выделены в два самостоятельных направления. Известно, что статьи по теме “Computer and information sciences” традиционно имеют более высокие показатели цитирования по сравнению с математическими статьями. Объединение этих двух направлений в РИНЦ могло привести, в частности, к тому, что в тематическом рейтинге по «Математике» журнал «Информатика и автоматизация» имеет лучшие показатели по сравнению с показателями классических математических журналов. Таким образом, от выбора классификатора зависят показатели эффективности журналов (позиции в тематических рейтингах).

В справочных материалах РИНЦ дана информация о применении методов нормализации показателей цитирования, учитывающих специфику тематических категорий на уровне журналов, при построении отдельных видов рейтингов журналов. Однако нормализация показателей цитирования на уровне отдельных статей, судя по справочным материалам РИНЦ, не применяется. При рассмотрении результатов рейтинга РИНЦ по импакт-фактору в направлении «Математика» было показано, что именно «непрофильные» статьи и цитирования могут незаслуженно вывести «непрофильные» журналы на первые позиции чужого рейтинга.

3. О КЛАССИФИКАТОРАХ

Как отмечают многие специалисты в области наукометрии, не существует идеальных классификаторов. В практической работе и интерпретации результатов работы сервисов библиографической базы приходится учитывать конкретные особенности выбранных систем тематической классификации.

Так, многие библиографические базы используют систему классификации Web of Science. Однако эта система имеет свои особенности, в частности, она не позволяет однозначно классифицировать содержание междисциплинарных журналов, таких как Science и Nature: эти и подобные журналы сложно отнести к одной или нескольким категориям. Следует иметь в виду, что тематические области в библиографической базе, как и категории Web of Science, как правило, предназначены для организации поиска информации, и их использование для иной цели может иметь нежелательные последствия.

Тема применения классификаторов является весьма чувствительной для разработчиков библиографических баз. В настоящее время существуют десятки и сотни классификаторов, привязанных к крупным библиотекам, известным библиографическим базам. Многие страны разрабатывают свои собственные национальные классификаторы научных областей. Существует проблема гармонизации различных систем тематической классификации.

В недавнем отчете компании Clarivate [10] представлен альтернативный подход к формированию референтных групп, который можно характеризовать как подход «снизу вверх». Предлагаемый подход альтернативен принятому сейчас подходу, основанному на использовании тематических классификаторов научных дисциплин (этот способ классификации можно назвать подходом «сверху вниз»).

Подход «снизу вверх» предлагает использовать в качестве меры тематической близости статей их библиографические связи: прямое цитирование, пересечения пристатейных библиографических списков (со-цитирование). В единый тематический кластер объединяются наиболее близкие статьи. Тематический кластер постоянно увеличивается по мере появления новых цитирующих статей. Есть и другой эффект: с появлением новых цитирований статья может оказаться включенной дополнительно в другой тематический кластер. Тематические кла-

стеры нижнего уровня объединяются в более крупные кластеры посредством предъявления более слабых требований к близости публикаций. Таким образом формируются кластеры более высокого уровня, отражающие более общие тематические категории.

Кластеры разных уровней получают названия (метки). Для кластеров нижнего уровня метка может формироваться автоматически, например, на основе наиболее значимых ключевых слов из статей кластера. Для кластеров среднего и высшего уровней метки (категории) могут назначаться экспертами в предметной области, привязываясь к привычным или удобным названиям, отражающим тематику статей, сгруппированных в кластере. Эксперты при этом могут опираться на характерные ключевые слова и привычные для ученых категории, в частности, категории Web of Science. Строящиеся таким образом тематические кластеры получили название “Citation Topics” («Темы цитирования») [10].

В декабре 2020 года такой новой системой классификации Citation Topics, разработанной совместно с ведущей научной группой из Центра научно-технологических исследований (CWTS) Лейденского университета (Нидерланды), был дополнен аналитический продукт InCites компании Clarivate. InCites — система, позволяющая получать разнообразные библиометрические показатели авторов, организаций, журналов, а также проводить многоаспектные наукометрические исследования. InCites предоставлял пользователю возможность отбора статей по классификаторам, построенным по принципу «сверху вниз». Теперь наряду с этим система включает классификатор Citation Topics [11], построенный по принципу формирования категорий «снизу вверх».

В данный момент алгоритмический классификатор Citation Topics включает 10 категорий высокого уровня, 326 категорий среднего уровня и 2444 категории (кластера) нижнего уровня. Алгоритм классификатора обработал более 60 млн статей, и более 50 млн были отнесены к тому или иному тематическому кластеру. Несмотря на принципиальное различие между методологиями классификации «снизу вверх» (кластеризация цитирований) и классификации «сверху вниз» (категории журналов), группы в этих структурах во многом совпадают.

Создание тематических классификаторов в описанной технологии «снизу-вверх» снимает некоторые проблемы, свойственные классификаторам, постро-

енным по принципу «сверху-вниз». В частности, технология тематической классификации «снизу вверх», на наш взгляд, дает более надежную объективную основу для формирования референтных групп с целью дальнейшего ранжирования или сравнения библиометрических показателей журналов.

В работе [12] представлены результаты исследования графа цитирования для журналов физико-математического направления, входящих в библиографическую базу Math-Net.Ru. Исследование проведено для 120 журналов, общий объем обработанных библиографических ссылок — около 96 тысяч. Построенный граф цитирования, идеологически близкий кластеризации Citation Topics, показал достаточно высокую плотность, свидетельствующую об активном взаимном цитировании статей, опубликованных в журналах. Граф цитирований распался на пять модулей (кластеров), которые, как оказалось, могут быть интерпретированы по направлениям исследований: фундаментальная математика, математическое моделирование, экспериментальная и теоретическая физика, дискретная математика и прикладная математика и компьютерные науки. Для графа цитирования был построен рейтинг журналов на основе аналога показателя PageRank [13]. Оказалось, что полученный рейтинг вполне согласуется с представлениями российских математиков об авторитете и популярности математических журналов, участвующих в рейтинге.

ЗАКЛЮЧЕНИЕ

Одна из основных проблем продуктивного построения рейтинга журналов в тематическом направлении связана с адекватным выделением референтной группы журналов. Тематика журналов такой группы определяется системой классификации научных направлений, используемой в библиографической базе. Недостаток существующих систем классификации, приводящий к появлению неубедительных рейтингов, заключается в их декларативном характере, основанном на классификации «сверху вниз». Декларируемое журналом тематическое направление нередко исходит из субъективных оценок и не отражает тематики публикуемых статей. Технология формирования тематических кластеров по принципу «снизу вверх» существенно точнее и объективнее структурирует множество тематически разнообразных журнальных статей. Тематическая близость статей определяется здесь посредством анализа близости библиографических

списков, выявления прямых ссылок и со-цитирований. Вычленение референтных групп журналов на основе подхода «снизу вверх» является более надежным фундаментом для построения тематических рейтингов, позволяя конструктивно учитывать мультидисциплинарность на уровне статей.

СПИСОК ЛИТЕРАТУРЫ

1. *Leydesdorff L., Wouters P., Bornmann L.* Professional and Citizen Bibliometrics: Complementarities and Ambivalences in the Development and Use of Indicators – A State-of-the-Art Report // *Scientometrics*. 2016. Vol. 109. No. 3. P. 2129–2150. <https://doi.org/10.1007/s11192-016-2150-8>

2. *Москалева О.В.* Научные публикации как средство коммуникации. В книге: *Руководство по наукометрии: индикаторы развития науки и технологии*. Акоев М.А., Маркусова В.А., Москалева О.В., Писляков В.В. Екатеринбург, 2021. С. 140–176. <https://doi.org/10.15826/B978-5-7996-3154-3>

3. *Akoev M., Markusova V., Moskaleva O., Pislyakov V.* Handbook on Scientometrics: Science and Technology Development Indicators, Second edition. 2021. 358 p. <https://doi.org/10.15826/B978-5-7996-3154-3>

4. *Hicks D., Wouters P., Waltman L. et al.* Bibliometrics: The Leiden Manifesto for research metrics. *Nature*. Vol. 520, 429–431 (2015). <https://doi.org/10.1038/520429a>

5. Российский индекс научного цитирования.
URL: https://www.elibrary.ru/project_risc.asp

6. Научная электронная библиотека eLibrary.ru.
URL: <https://www.elibrary.ru/defaultx.asp>

7. *Общероссийский портал Math-Net.Ru*. URL: <http://www.mathnet.ru/>

8. *Полилова Т.А.* Рейтинги журналов в РИНЦ как инструменты анализа и влияния // *Препринты ИПМ им. М.В. Келдыша*. 2021. № 40. 35 с. <https://doi.org/10.20948/prepr-2021-40>

9. *Vinkler P.* Evaluation of publications by the part-set method // *Scientometrics*. 2021. V. 126. P. 2737–2757. <https://doi.org/10.1007/s11192-020-03841-7>

10. *Шомшор М. (Martin Szomszor), Адамс Д. (Jonathan Adams), Пендлбери Д. (David A. Pendlebury), Роджерс Г. (Gordon Rogers)*. Классификация

данных: как делать осознанный выбор, ведущий к желаемым результатам. Отчет о международном исследовании Института научной информации (ISI).

URL: https://img06.en25.com/Web/ClarivateAnalytics/%7B3f970f80-b453-451f-80a2-8d1792213e83%7D_Clarivate_ISI_Data_categorization_report_RU.pdf

11. *Potter I.* Introducing Citation Topics in InCites – Clarivate. Dec 3, 2020. URL: <https://clarivate.com/blog/introducing-citation-topics/>

12. *Печников А.А., Чебуков Д.Е.* Структура графа цитирования журналов Math-Net.Ru // Научный сервис в сети Интернет: труды XXIII Всероссийской научной конференции (20–24 сентября 2021 г., онлайн). М.: ИПМ им. М.В. Келдыша, 2021. <https://doi.org/10.20948/abrau-2021-2>

13. *PageRank.* Wikipedia, the free encyclopedia.
URL: <https://en.wikipedia.org/wiki/PageRank>

THE RATING OF THE JOURNAL IN THE BIBLIOGRAPHIC DATABASE

M. M. Gorbunov-Posadov¹ [0000-0002-7044-8287], **T. A. Polilova**² [0000-0003-4628-3205]

^{1, 2}Keldysh Institute of Applied Mathematics, Miusskaya sq., 4, Moscow, 125047, Russia

¹gorbunov@keldysh.ru, ²polilova@keldysh.ru

Abstract

The tool for building ratings of scientific journals is one of the popular services of bibliographic databases. The task of building a rating is usually divided into two main subtasks: determining the reference group of journals and calculating the rating indicator for journals of this reference group. Practice shows that for the correct comparison of journals, a necessary condition is to limit the reference group to exclusively journals of a certain subject. In the case of methodological errors made at the stage of selecting a reference group, the values of the journal index in the rating may differ greatly from the expected ones.

For example, in the ranking of journals in the Russian Science Citation Index (RSCI) according to the two-year impact factor in the thematic area “Mathematics”, classical fundamental mathematical journals, contrary to expectations, do not reach

the first positions of the rating. The first positions were taken by journals for which mathematics is not the dominant profile discipline. Analysis of statistical data on the subject of published articles and citations in journals that occupy leading positions in the RSCI rating shows that the multidisciplinary nature of these journals significantly influenced the rating indicators.

The noted misunderstanding leads to the idea that in this case, not all the articles of the journal should have been involved in the calculation of the rating, but only those related to this thematic area. At the same time, the existing scheme of thematic classification of directions also raises questions. The "bottom-up" classification, which is gaining popularity and works on a representative array of articles, seems to be more promising. Here thematic clusters are isolated on the basis of the concept of proximity of articles, interpreted as the proximity of their bibliographic links. And further, the thematic affiliation of the article is not assigned by the volitional decision of the author or the editorial board, but is strictly formally calculated on the basis of its bibliographic list.

Keywords: *scientific publication, citation, rating of journals, thematic classification, impact factor, multidisciplinary, bibliographic reference, co-citation, bottom-up classification, thematic clustering, Citation Topics.*

REFERENCES

1. Leydesdorff L., Wouters P., Bornmann L. Professional and Citizen Bibliometrics: Complementarities and Ambivalences in the Development and Use of Indicators – A State-of-the-Art Report // *Scientometrics*. 2016. Vol. 109, No. 3. P. 2129–2150. <https://doi.org/10.1007/s11192-016-2150-8>

2. Moskaleva O.V. Nauchnye publikatsii kak sredstvo kommunikatsii. V knige: *Rukovodstvo po naukometrii: indikatory razvitiia nauki i tekhnologii*. Akoev M.A., Markusova V.A., Moskaleva O.V., Pisiakov V.V. Ekaterinburg, 2021. S. 140–176. <https://doi.org/10.15826/B978-5-7996-3154-3>.

3. Akoev M., Markusova V., Moskaleva O., Pisiakov V. *Handbook on Scientometrics: Science and Technology Development Indicators*, Second edition. 2021. 358 p. <https://doi.org/10.15826/B978-5-7996-3154-3>.

4. *Hicks D., Wouters P., Waltman L. et al.* Bibliometrics: The Leiden Manifesto for research metrics // *Nature*. 2015. V. 520. P. 429–431.

<https://doi.org/10.1038/520429a>

5. Rossiiskii indeks nauchnogo tsitirovaniia.

URL: https://www.elibrary.ru/project_risc.asp

6. Nauchnaia elektronnaia biblioteka eLibrary.ru.

URL: <https://www.elibrary.ru/defaultx.asp>

7. Obshcherossiiskii portal Math-Net.Ru. URL: <http://www.mathnet.ru/>

8. *Polilova T.A.* Reitingi zhurnalov v RINTs kak instrumenty analiza i vliianiia // *Preprinty IPM im. M.V. Keldysha*. 2021. № 40. 35 s. <https://doi.org/10.20948/prepr-2021-40>

9. *Vinkler P.* Evaluation of publications by the part-set method // *Scientometrics*. 2021. Vol. 126. P. 2737–2757. <https://doi.org/10.1007/s11192-020-03841-7>

10. *Shomshor M. (Martin Szomszor), Adams D. (Jonathan Adams), Pendlberi D. (David A. Pendlebury), Rodzhers G. (Gordon Rogers).* Klassifikatsiia dannykh: kak delat osoznannyi vybor, vedushchii k zhelaemym rezultatam. Otchet o mezhdunarodnom issledovanii Instituta nauchnoi informatsii (ISI).

URL: https://img06.en25.com/Web/ClarivateAnalytics/%7B3f970f80-b453-451f-80a2-8d1792213e83%7D_Clarivate_ISI_Data_categorization_report_RU.pdf

11. *Potter I.* Introducing Citation Topics in InCites – Clarivate. Dec 3, 2020. URL: <https://clarivate.com/blog/introducing-citation-topics/>

12. *Pechnikov A.A., Chebukov D.E.* Struktura grafa tsitirovaniia zhurnalov Math-Net.Ru // *Nauchnyi servis v seti Internet: trudy XXIII Vserossiiskoi nauchnoi konferentsii (20–24 sentiabria 2021 g., onlain)*. M.: IPM im. M.V. Keldysha, 2021.

<https://doi.org/10.20948/abrau-2021-2>

13. *PageRank*. Wikipedia, the free encyclopedia.

URL: <https://en.wikipedia.org/wiki/PageRank>

СВЕДЕНИЯ ОБ АВТОРАХ



ГОРБУНОВ-ПОСАДОВ Михаил Михайлович – главный научный сотрудник Института прикладной математики им. М.В. Келдыша РАН, доктор физико-математических наук;

Mikhail Mikhailovich GORBUNOV-POSADOV – Keldysh Institute of Applied Mathematics, chief researcher

email: gorbunov@keldysh.ru

ORCID: 0000-0002-7044-8287



ПОЛИЛОВА Татьяна Алексеевна – старший научный сотрудник Института прикладной математики им. М.В. Келдыша РАН, доктор физико-математических наук, лауреат Премии Президента РФ в области образования;

Tatyana Alekseevna POLILOVA – senior researcher of the Keldysh Institute of Applied Mathematics Russian Academy of Sciences.

email: polilova@keldysh.ru.

ORCID: 0000-0003-4628-3205

Материал поступил в редакцию 21 октября 2021 года