

КАЧЕСТВЕННЫЙ И КОЛИЧЕСТВЕННЫЙ АНАЛИЗ ПУБЛИКАЦИОННОЙ АКТИВНОСТИ РОССИЙСКИХ НАУЧНЫХ УЧРЕЖДЕНИЙ В ПРЕДМЕТНОЙ ОБЛАСТИ «ПРИКЛАДНАЯ ХИМИЯ»

© В. В. Королева¹, О. В. Иванов¹, А. А. Ведягин^{1,2*}, А. С. Лядов^{1,3},
А. В. Леонидов¹, А. В. Колобов¹

¹ Физический институт им. П. Н. Лебедева РАН,
119991, г. Москва, Ленинский пр., д. 53

² Институт катализа им. Г. К. Борескова СО РАН,
630090, г. Новосибирск, пр. Академика Лаврентьева, д. 5

³ Институт нефтехимического синтеза им. А. В. Топчиева РАН,
119991, г. Москва, Ленинский пр., д. 29

* E-mail: vedyagin@catalysis.ru

Поступила в Редакцию 21 августа 2020 г.

После доработки 21 августа 2020 г.

Принята к публикации 25 августа 2020 г.

Статья посвящена качественному и количественному анализу публикационной активности академических организаций химического профиля, подведомственных Министерству науки и высшего образования РФ, за период 2015–2019 гг. в предметной области «Прикладная химия». В качестве источника информации была использована база данных Web of Science, анализ проведен с использованием интеллектуальной системы анализа и прогнозирования научной деятельности «Зеркало», разработанной Физическим институтом им. П. Н. Лебедева РАН. Выявлены наиболее значимые ключевые слова, которые в значительной степени характеризуют перспективные направления в области прикладной химии. Методом фракционного счета с дифференциацией вкладов от статей проведен анализ публикационной результативности научных организаций и определены периоды публикационной активности. Проведен геолокационный анализ публикаций на основе данных аффилиций, указанных в статье, что позволило установить наличие устойчивых международных связей у трети авторских коллективов.

Ключевые слова: результативность научной деятельности; фракционный счет; комплексной балл публикационной результативности (КБПР); база данных Web of Science

DOI: 10.31857/S0044461820090030

Введение

В последние годы в результате интенсивного развития средств математической статистики, методов оптимизации, новых алгоритмов работы с данными в цифровой форме, а также расширения областей применения теории графов существенно возросла

популярность и востребованность методов искусственного интеллекта для анализа массивов данных, включая «большие данные» (Big Data), в различных сферах человеческой деятельности. Одним из наиболее часто реализуемых приложений данных методов при оценке научной деятельности является анализ публикационной активности в той или иной темати-

ческой области [1–5]. Подобный анализ может быть использован как для определения новых тенденций развития науки и техники, установления факта зарождения новых научных направлений, так и для оценки эффективности деятельности организаций, исследовательских коллективов или отдельно взятых персоналий [6].

Из существующих в научном мировом пространстве баз данных наибольшей популярностью пользуются такие платформы, как Web of Science и Scopus [7, 8], которые объединяют реферативные базы публикаций в научных периодических изданиях, включая базы по взаимному цитированию публикаций. Информация, собранная в этих базах данных, является наиболее полной с точки зрения оценки индекса цитирования по отношению к альтернативным базам данных [Российский индекс научного цитирования (РИНЦ), Google Scholar]. К дополнительным факторам, привлекающим внимание аналитиков, можно отнести возможность реализации механизма обмена данными с другими базами данных, что существенно упрощает процессы роботизированного извлечения информации для конкретных целей.

Цель работы — анализ качественных и количественных аспектов публикационной активности научных организаций химического профиля, находящихся в Российской Федерации (35 организаций), по направлению «Прикладная химия» на основе базы данных Web of Science с использованием интеллектуальной системы анализа и прогнозирования научной деятельности «Зеркало».

Методологическая часть

Тематический анализ и фракционный счет публикационной активности за период 2015–2019 гг. был проведен для 35 научных организаций химического профиля на основе базы данных Web of Science с использованием программно-аналитической системы «Зеркало», разработанной в Физическом институте им. П. Н. Лебедева РАН. При формировании первичной базы данных публикаций и ее анализе был использован фильтр по категории Web of Science «Chemistry, Applied», соответствующий предметной области «Прикладная химия». Тематический анализ проводили на основе полученного массива ключевых слов. В основе фракционного счета был использован комплексный балл публикационной результативности (КБПР), который рассчитывали по формуле*

$$\text{КБПР}_{\text{num}} = \sum_{k=1}^n T_{\text{num}}^k, \quad (1)$$

где n — полное число строк Организации в публикационном отчете Минобрнауки России, T_{num}^k — балл k -той строки публикационного отчета для Организации с номером num.

Балл k -той строки был рассчитан по формуле

$$T_{\text{num}}^k = K_m \frac{1}{N^m} \frac{1}{a^m} A_{\text{num}}, \quad (2)$$

где m — уникальная статья; N^m — число авторов в статье m ; a^m — количество аффилиаций автора в статье m ; $A_{\text{num}} = 1$, если автор статьи m указал аффилиацию num; $A_{\text{num}} = 0$, если автор статьи m не указал аффилиацию num; K_m — коэффициент качества статьи/журнала (см. таблицу).

Обсуждение результатов

Всего в период 2015–2019 гг. организациями химического профиля было опубликовано 4450 статей в предметной области «Прикладная химия». Данная область включает научные работы, направленные на решение проблем направленного синтеза промышленно важных продуктов. Исторически к таким процессам относятся производство серной и азотной кислот, аммиака, ацетилена, спиртов, альдегидов, уксусной кислоты, полимеров и химических волокон, минеральных удобрений, силикатных материалов, металлов и т. д. Кроме того, следует отметить процессы химической переработки жидкого и твердого топлива, а также производство и переработку углеводородных газов. Немаловажную роль играет основной органический синтез и производство продукции тонкого органического синтеза.

В ходе анализа статей, опубликованных организациями химического профиля в период 2015–2019 гг., был сформирован массив ключевых слов, отражающий специфику проводимых исследований. Общее количество уникальных ключевых слов составило 581 (рис. 1). Химический термин «производные» (derivatives) является абсолютным лидером по количеству статей, в которых он упоминается в качестве ключевого слова. Термин «адсорбция» (adsorption) встречается в меньшем количестве статей, но лидирует по значению КБПР, количество статей для которого более чем в 3 раза ниже. Данный факт сви-

* Методика расчета качественного показателя государственного задания «Комплексный балл публикационной

результативности» для научных организаций, подведомственных Министерству науки и высшего образования Российской Федерации (утв. зам. министра науки и высшего образования РФ 30.12.2019).

Коэффициенты качества статей в зависимости от журнала, в котором они опубликованы

Квартиль журнала по данным WoS				Публикации в изданиях без квартиля, но входящие в WoS Core Collection	Публикации в изданиях, индексируемых в Scopus и неиндексируемых в WoS	Публикации в журналах из RSCI Web of Science, неиндексируемых в Core Collection Web of Science и Scopus (по данным РИНЦ)	Публикации в журналах списка ВАК, не входящих в вышеперечисленные пункты (по данным РИНЦ)	Монографии, зарегистрированные в Российской книжной палате
1	2	3	4					
19.7	7.3	2.7	1	1	1	0.75	0.5	1

детельствует о публикации материалов по адсорбции в более высокорейтинговых журналах. Третье место занимает термин «комплексные соединения» (complexes). Далее следуют термины «наночастицы» (nanoparticles), «окисление» (oxidation), «механизм» (mechanism) и «эксплуатационные показатели» (performance). Ввиду относительно небольшого числа публикаций при достаточно высоком значении КБПР можно предположить, что данные работы также занимают высокое место в рейтинге исследуемой предметной области. Исследование окислительных процессов и их механизмов, а также уникальных физико-химических и каталитических свойств наночастиц, действительно, является весьма актуальным направлением [9]. Замыкают рейтинг такие термины, как «кинетика» (kinetics), «поведение» (behavior) и «кристаллическая структура» (crystal structure). Эти три термина отражают важные направления фундаментальных исследований, которые в ближайшее

время могут быть внедрены в промышленность. Исследования кинетики необходимы как для установления детального механизма процесса, так и для расчета параметров реакторов для проведения этих процессов. Поведение материалов в реакционных условиях позволяет определить область их потенциального применения. Кристаллическая структура в свою очередь во многом определяет свойства материалов. Таким образом, перечисленные ключевые слова наиболее точно характеризуют изучаемую предметную область «Прикладная химия».

Одной из проблем количественного анализа публикационной результативности организаций является учет всех публикаций, в которых хотя бы у одного из авторов в качестве одной из аффилиаций указана данная конкретная организация. Так, например, если авторы из трех организаций опубликуют совместно одну статью и каждой организации в зачет пойдет одна статья, то при суммировании будет получено значение 3, что позволяет одной статьей отчитываться трем организациям, например, по государственному заданию. Данная проблема может быть решена при использовании фракционного счета, когда в зачет идут не целые, а дробные единицы, учитывающие доли авторов данной организации в общем числе авторов, а также аффилиационную долю каждого из авторов. Проведенный анализ показал, что в период 2015–2018 гг. происходил монотонный рост числа публикаций, а в 2019 г. наблюдался спад до уровня 2016 г. (рис. 2). При анализе значений КБПР, учитывающих качество публикаций (табл. 1), характер зависимости несколько изменялся. В этом случае монотонный рост наблюдался только до 2017 г., после чего следовал резкий всплеск в 2018 г. Можно заключить, что в 2018 г. было опубликовано большее количество статей в более высокорейтинговых журналах. Следует также отметить, что значение КБПР для 2019 г. превосходит значения для периода 2015–2017 гг., что косвенно свидетельствует о более качественных публикациях. Для вычленения качественной компоненты из количественно-качественного показателя КБПР удобно использовать удельный показатель KBP/N , где N — общее чис-

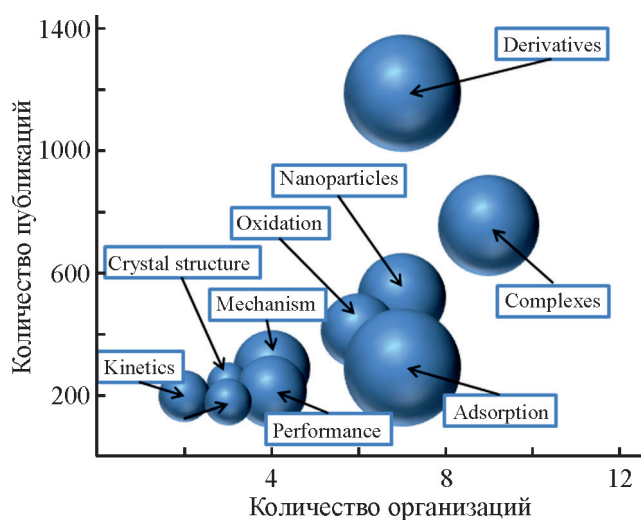


Рис. 1. Наиболее часто употребляемые ключевые слова в статьях, опубликованных в предметной области «Прикладная химия» [по оси абсцисс отложено количество организаций (из 35), в статьях которых присутствуют данные ключевые слова, по оси ординат — общее количество статей с этими ключевыми словами; размер пузырьков соответствует суммарному значению КБПР этих же статей].

ло статей, опубликованных в анализируемый год. При использовании удельного показателя ситуация изменяется кардинально (рис. 2). 2017 и 2018 г. характеризуются наименьшими значениями удельного показателя, а в 2019 г. достигается столь же высокое значение, что наблюдали в 2015 г. Таким образом, методика фракционного счета с дифференциацией вкладов от статей, опубликованных в журналах из различных квартилей, позволяет проводить более глубокий анализ публикационной результативности научных организаций.

Распределение статей, опубликованных в период 2015–2019 гг., по квартилям базы данных WoS (рис. 3, а) показывает, что по количеству статей, опубликованных в предметной области «Прикладная химия», вклад журналов первого (Q1) и четвертого (Q4) квартилей примерно одинаков и составляет 39 и 41%. Основная доля статей Q4 была опубликована в Журнале прикладной химии. Далее следуют журналы Q3 — 19%. На журналы Q3 приходится порядка 1%, а доля остальных журналов, не входящих в квартили WoS, не превышает 1%. Это же распределение, построенное в значениях КБПР, выглядит иначе (рис. 3, б): 39% статей Q1 дают вклад 77%, 19% статей Q2 — 17%, а вклад 41% статей Q4 составляет всего лишь 6%.

Интеллектуальная система «Зеркало» позволяет также по идентификаторам авторов оценить продолжительность публикационной активности по годам (рис. 4). Организации отсортированы снизу вверх по увеличению доли авторов, активных в течение всех

пяти лет. Весомый вклад именно этих авторов позволяет говорить о наличии в организациях научных школ или активно работающих научных коллективов. Авторы, задействованные в течение лишь одного года, есть во всех организациях, и их доля варьируется от 22 до 71%. В среднем около 40% приходится на авторов, проявляющих публикационную активность в течение 2–4 лет. Для последней категории авторов, публикующихся в течение пяти лет, разброс составляет от 0 до 27%. Таким образом, можно сделать вывод, что существуют организации, в которых основной вклад вносят сотрудники, активные в течение 1–2 лет (рис. 4, № 1), при том что для большинства организаций авторы всех категорий вовлечены в публикационный процесс более равномерно (рис. 4, № 10–35).

Еще одна возможность системы «Зеркало» заключается в определении степени локализации авторов исходя из анализа данных их аффилиаций. Данная опция позволяет разделить весь анализируемый массив публикаций на три типа. В публикациях первого типа все авторы аффилированы только с российскими организациями. Второй тип — это смешанные публикации, в которых у части авторов есть зарубежные аффилиации, но другая часть аффилирована только к российским учреждениям. Все авторы публикаций третьего типа имеют как минимум одну аффилиацию к зарубежным организациям.

Все анализируемые организации были отсортированы снизу вверх по убыванию доли публикаций первого типа (рис. 5). Выявлено, что всего лишь

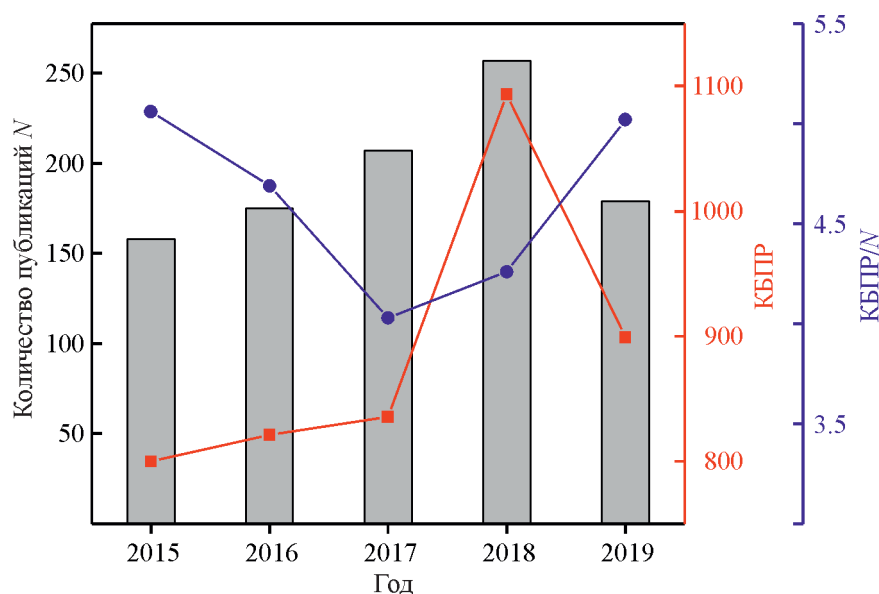


Рис. 2. Динамика изменения общего числа публикаций, а также абсолютных и удельных значений КБПР в период 2015–2019 гг. в предметной области «Прикладная химия» для 35 научных организаций химического профиля, находящихся в Российской Федерации.

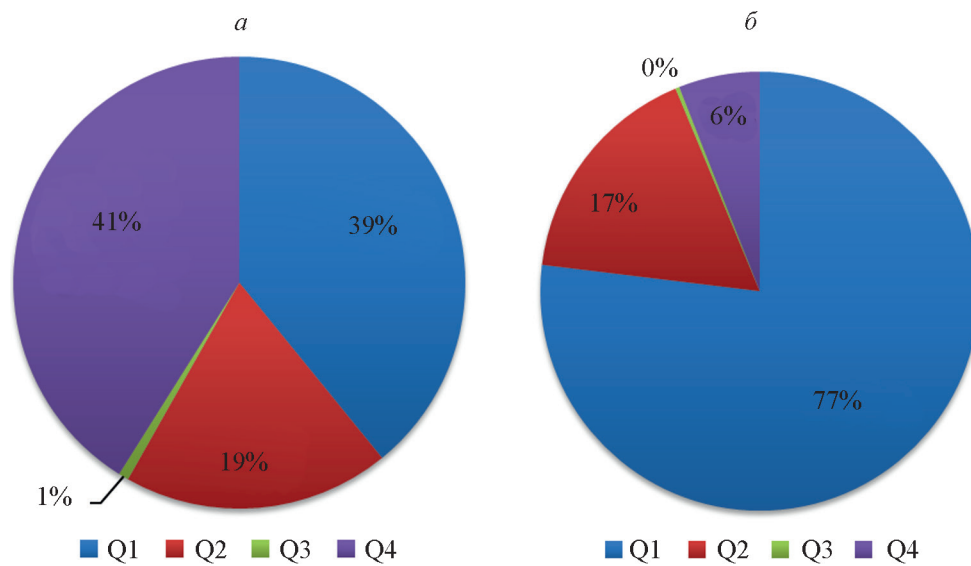


Рис. 3. Распределение статей, опубликованных в период 2015–2019 гг., по квартилям журналов (WoS): *a* — по количеству статей; *б* — по значению КБПР, рассчитанному для этих статей.

в трех организациях имеются публикации третьего типа, т. е. сотрудник данной организации имеет вторую зарубежную аффилиацию. В среднем вклад публикаций первого типа превышает 60%, хотя для

13 организаций доля смешанных публикаций (типа II) составляет 45–65%. Наличие подобных публикаций свидетельствует прежде всего об устойчивых международных связях авторских коллективов. Как

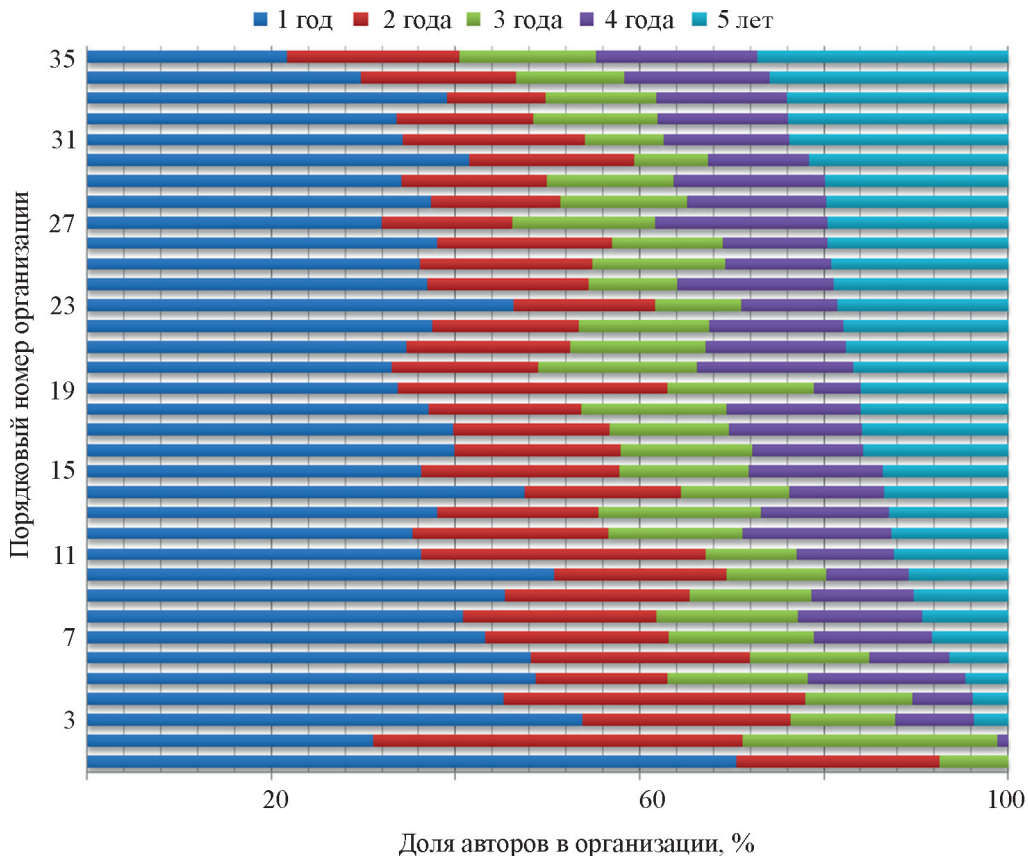


Рис. 4. Распределение авторов по длительности публикационной активности в 2015–2019 гг. для 35 анализируемых организаций.

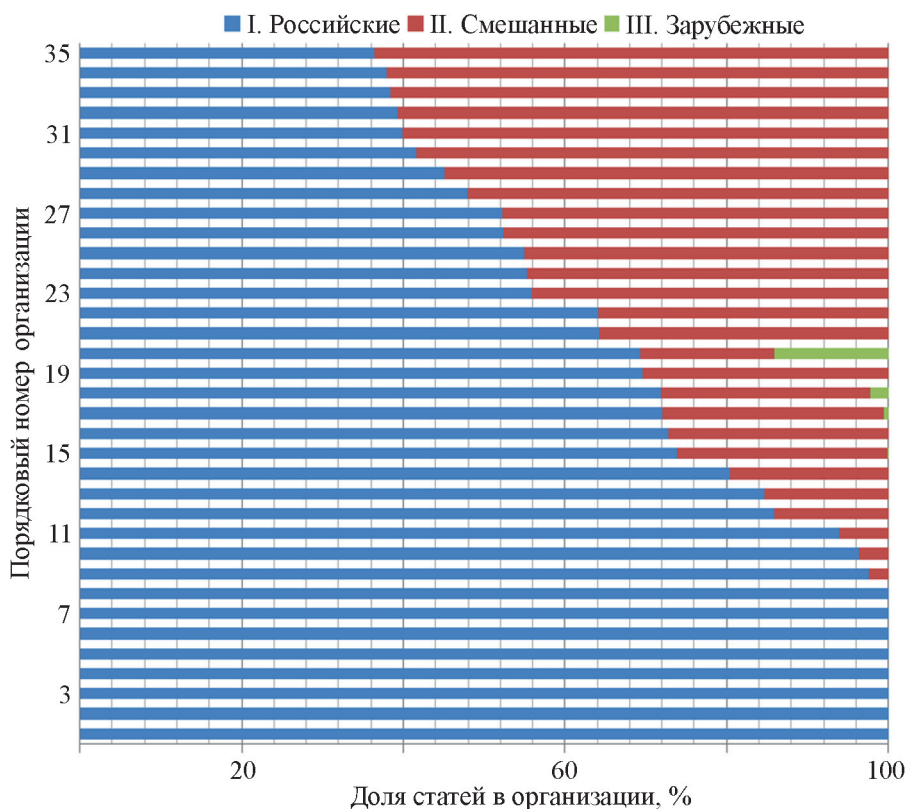


Рис. 5. Распределение статей, опубликованных в 2015–2019 гг., по типу аффилиаций авторов для 35 анализируемых организаций.

I — все авторы имеют российскую аффилиацию, II — часть авторов имеет зарубежную аффилиацию, III — все авторы имеют зарубежную аффилиацию.

правило, финансовая поддержка таких исследовательских работ осуществляется за счет зарубежных фондов или в рамках межправительственных соглашений.

Заключение

В ходе работы на примере предметной области «Прикладная химия» продемонстрированы возможности интеллектуальной системы анализа и прогнозирования научной деятельности «Зеркало» с использованием базы данных Web of Science. Проведен качественный и количественный анализ публикационной активности 35 научных организаций химического профиля Российской Федерации за период 2015–2019 гг. Выявлены наиболее значимые ключевые слова, которые в значительной степени характеризуют перспективные направления в области прикладной химии. Показано, что методика фракционного счета с дифференциацией вкладов от статей, опубликованных в журналах из различных квартилей, позволяет проводить более глубокий анализ публикационной результативности научных организаций.

Установлено, что в целом научные организации характеризуются малым временем публикационной активности (в среднем 1–2 года) по исследуемому направлению, что является недостаточным для формирования научных школ или активно работающих научных коллективов. Геолокализационный анализ публикаций по аффилиационной привязке авторов позволил установить наличие устойчивых международных связей примерно у трети авторских коллективов.

Финансирование работы

Работа выполнена при финансовой поддержке Министерства науки и высшего образования Российской Федерации, соглашение о предоставлении субсидии № 05.601.21.0020 от 8 ноября 2019 г. (уникальный идентификатор соглашения RFMEFI60119X0020) «Исследование механизмов адаптивного формирования кадрового потенциала для проведения разномасштабных программ исследований по приоритетным направлениям научно-технологического развития Российской Федерации».

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов, требующего раскрытия в данной статье.

Информация об авторах

Королева Валентина Владимировна,
ORCID: <http://orcid.org/0000-0003-2363-659X>
Иванов Олег Викторович, к.ф.-м.н.,
ORCID: <http://orcid.org/0000-0002-7658-305X>
Ведагин Алексей Анатольевич, к.х.н., доцент,
ORCID: <http://orcid.org/0000-0002-6930-936X>
Лядов Антон Сергеевич, к.х.н.,
ORCID: <http://orcid.org/0000-0001-9969-7706>
Леонидов Андрей Владимирович, д.ф.-м.н.,
ORCID: <http://orcid.org/0000-0002-6714-6261>
Колобов Андрей Владимирович, к.ф.-м.н.,
ORCID: <http://orcid.org/0000-0002-1713-8036>

Информация о вкладе авторов

В. В. Королева — проведение анализа на основе базы данных Web of Science с использованием программно-аналитической системы «Зеркало»; О. В. Иванов — постановка задачи и разработка методологии проведения исследования; А. А. Ведагин — анализ и обсуждение полученных результатов, написание текста статьи; А. С. Лядов — анализ и обсуждение полученных результатов, подготовка статьи к публикации; А. В. Леонидов — оптимизация работы программно-аналитической системы «Зеркало» под задачи проведенного исследования; А. В. Колобов — расчет и анализ данных, основанных на комплексном балле публикационной результативности.

Список литературы

- [1] Zhang J. Z., Cenci J., Becue V., Koutra S., Ioakimidis C. S. Recent evolution of research on industrial heritage in Western Europe and China based on bibliometric analysis // *Sustainability*. 2020. V. 12. N 13. ID 5348. <https://doi.org/10.3390/su12135348>
- [2] Yue T., Liu H. W., Long R. Y., Chen H., Gan X., Liu J. L. Research trends and hotspots related to global carbon footprint based on bibliometric analysis: 2007–2018 // *Environ. Sci. Pollut. Res.* 2020. V. 27. N 15. P. 17671–17691. <https://doi.org/10.1007/s11356-020-08158-9>
- [3] Usman M., Ho Y. S. A bibliometric study of the Fenton oxidation for soil and water remediation // *J. Environ. Manage.* 2020. V. 270. ID 110886. <https://doi.org/10.1016/j.jenvman.2020.110886>
- [4] Sweileh W. M. Bibliometric analysis of scientific publications on «sustainable development goals» with emphasis on «good health and well-being» goal (2015–2019) // *Glob. Health.* 2020. V. 16. N 1. ID 68. <https://doi.org/10.1186/s12992-020-00602-2>
- [5] Liu Z. Q., Yang J. Y., Zhang J. E., Xiang H. M., Wei H. A bibliometric analysis of research on acid rain // *Sustainability*. 2019. V. 11. N 11. ID 3077. <https://doi.org/10.3390/su11113077>
- [6] Зибарева И. В., Ильина Л. Ю., Альперин Б. Л., Ведагин А. А. Институт катализа СО РАН в зеркале наукометрии // *Вестн. РАН*. 2019. Т. 89. № 3. С. 243–254. <https://doi.org/10.31857/S0869-5873893243-254> [Zibareva I. V., Ilina L. Y., Alperin B. L., Vedyagin A. A. The scientometric profile of Boreskov Institute of Catalysis // *Herald Russ. Acad. Sci.* 2019. V. 89. N 3. P. 259–270. <https://doi.org/10.1134/S1019331619030109>].
- [7] Martín-Martín A., Orduna-Malea E., Thelwall M., Delgado López-Cózar E. Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories // *J. Informetr.* 2018. V. 12. N 4. P. 1160–1177. <https://doi.org/10.1016/j.joi.2018.09.002>
- [8] Zhu J., Liu W. A tale of two databases: The use of Web of Science and Scopus in academic papers // *Scientometrics*. 2020. V. 123. N 1. P. 321–335. <https://doi.org/10.1007/s11192-020-03387-8>
- [9] Zibareva I. V., Ilina L. Y., Vedyagin A. A. Catalysis by nanoparticles: The main features and trends // *React. Kinet., Mech. Catal.* 2019. V. 127. N 1. P. 19–24. <https://doi.org/10.1007/s11144-019-01552-6>