

УДК 004.8:005.94

НАУЧНО-ПУБЛИКАЦИОННАЯ АКТИВНОСТЬ В СФЕРЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА (II КВАРТАЛ 2023 ГОДА)

(Материал предоставлен департаментом по исследовательским центрам в сфере искусственного интеллекта Аналитического центра при Правительстве Российской Федерации совместно с Национальным центром развития искусственного интеллекта при Правительстве Российской Федерации)

SCIENTIFIC PUBLICATION ACTIVITY IN THE FIELD OF ARTIFICIAL INTELLIGENCE – Q2 2023

(The material was provided by the Department for Research Centers in the Field of Artificial Intelligence of the Analytical Center for the Government of the Russian Federation together with the National Centre for AI Development for the Government of the Russian Federation)

АСТАХОВ

Сергей Владимирович,

кандидат физико-математических наук, заместитель руководителя департамента по исследовательским центрам в сфере искусственного интеллекта Аналитического центра при Правительстве Российской Федерации

НАКВАСИН

Сергей Юрьевич,

директор Национального центра развития искусственного интеллекта при Правительстве Российской Федерации

Sergey Astakhov,

Ph.D candidate in Physics & Mathematics, Deputy Head of the Department for Research Centers in the Field of Artificial Intelligence of the Analytical Center for the Government of the Russian Federation

Sergei Nakvasin,

Director of the National Centre for AI Development for the Government of the Russian Federation

Аннотация: лидером по научно-публикационной активности в сфере искусственного интеллекта за второй квартал 2023 года является Китай, сохранивший свое место в рейтинге в течение последних четырех кварталов. При этом Россия находится на 16-м месте, демонстрируя укрепление своей позиции по отношению к прошлому кварталу, где Россия находилась на 19-м месте. Наиболее популярной темой по числу публикаций во втором квартале 2023 года является общая медицина. При этом самой востребованной темой по числу цитирований является программное обеспечение. В целом перечни из пяти самых популярных и самых востребованных тем соответствуют друг другу.

ABSTRACT: THE LEADER IN SCIENTIFIC PUBLICATION ACTIVITY IN THE FIELD OF ARTIFICIAL INTELLIGENCE FOR THE SECOND QUARTER OF 2023 IS CHINA, WHICH HAS MAINTAINED ITS PLACE IN THE RANKING FOR THE LAST FOUR QUARTERS. AT THE SAME TIME, RUSSIA IS IN 16TH PLACE, DEMONSTRATING THE STRENGTHENING OF ITS POSITION IN RELATION TO THE LAST QUARTER, WHERE RUSSIA WAS IN 19TH PLACE.

THE MOST POPULAR TOPIC BY THE NUMBER OF PUBLICATIONS IN THE SECOND QUARTER OF 2023 IS GENERAL MEDICINE. AT THE SAME TIME, THE MOST POPULAR TOPIC IN TERMS OF THE NUMBER OF CITATIONS IS SOFTWARE. IN GENERAL, THE LISTS OF THE FIVE MOST POPULAR AND MOST DEMANDED TOPICS CORRESPOND TO EACH OTHER.

По числу публикаций в наиболее популярных темах лидируют Китай (наиболее активная организация – Китайская академия наук) и США (наиболее активные организации – университеты Нортирстен, Гарвард и Эй-энд-эм). Эти же страны и организации сохраняют свое лидерство в наиболее востребованных темах. Россия самые высокие позиции занимает в области общей медицины и электроники и электротехники.

Наиболее авторитетным научным журналом, опубликовавшим статьи в сфере искусственного интеллекта за второй квартал 2023 года, является Information Sciences, отнесенный к квартилю цитируемости 1 и характеризующийся импакт-фактором 2,29 (SJR). Основная тематика журнала: информатика, изучение искусственного интеллекта и технологий поддержки принятия решений.

Доля статей с наиболее значимыми результатами составляет 0,7% от общего числа публикаций.

Лидером по публикациям в наиболее авторитетных журналах за второй квартал 2023 года является Китай.

Коммерческие компании публикуют результаты научных исследований по следующим темам: прикладная информатика, программное обеспечение, общая физика и математика, искусственный интеллект, общая медицина, компьютерные сети.

Ключевые слова: искусственный интеллект, публикационная активность, наукометрия, управление знаниями.

ACCORDING TO THE NUMBER OF PUBLICATIONS IN THE MOST POPULAR TOPICS, CHINA (THE MOST ACTIVE ORGANIZATION IS THE CHINESE ACADEMY OF SCIENCES) AND THE USA (THE MOST ACTIVE ORGANIZATIONS ARE THE UNIVERSITIES OF NORTIRSTEN, HARVARD AND A&M) ARE LEADING. THESE SAME COUNTRIES AND ORGANIZATIONS RETAIN THEIR LEADERSHIP IN THE MOST POPULAR TOPICS. RUSSIA OCCUPIES THE HIGHEST POSITIONS IN THE FIELD OF GENERAL MEDICINE AND ELECTRONICS AND ELECTRICAL ENGINEERING.

THE MOST AUTHORITATIVE SCIENTIFIC JOURNAL THAT HAS PUBLISHED ARTICLES IN THE FIELD OF ARTIFICIAL INTELLIGENCE FOR THE SECOND QUARTER OF 2023 IS INFORMATION SCIENCES, ASSIGNED TO THE CITATION QUARTILE 1 AND CHARACTERIZED BY AN IMPACT FACTOR OF 2.29 (SJR). THE MAIN TOPICS OF THE JOURNAL ARE COMPUTERSCIENCE, THE STUDY OF ARTIFICIAL INTELLIGENCE, THE STUDY OF DECISION SUPPORT TECHNOLOGIES.

THE SHARE OF ARTICLES WITH THE MOST SIGNIFICANT RESULTS IS 0.7% OF THE TOTAL NUMBER OF PUBLICATIONS.

THE LEADER IN PUBLICATIONS IN THE MOST AUTHORITATIVE JOURNALS FOR THE SECOND QUARTER OF 2023 IS CHINA.

COMMERCIAL COMPANIES PUBLISH THE RESULTS OF SCIENTIFIC RESEARCH ON THE FOLLOWING TOPICS: APPLIED COMPUTER SCIENCE, SOFTWARE, GENERAL PHYSICS AND MATHEMATICS, ARTIFICIAL INTELLIGENCE, GENERAL MEDICINE, COMPUTER NETWORKS.

Keywords: artificial intelligence, publication activities, scientometrics, knowledge management.

СИТУАЦИЯ В МИРЕ

По данным ресурса Lens.org, за второй квартал 2023 года было опубликовано 119 523 статьи. На рисунках 1 и 2 приведены графики квартальной динамики числа публикаций за последний год и за прошедшие пять лет соответственно.

Результаты, представленные на рисунках 1 и 2, позволяют сделать вывод о том, что мировая ежеквартальная научно-публикационная активность в сфере искусственного интеллекта показала спад с четвертого квартала 2022 года, при этом роста в течение текущего года также не наблюдается, число публикаций в текущий

период сопоставимо со значениями прошлого года за тот же отрезок времени, учитывая продолжающийся процесс индексации публикаций в рассматриваемом периоде. Следует отметить, что в конце прошлого квартала число проиндексированных публикаций составляло 182 814.

В связи с завершением индексации статей за первый квартал 2023 года их значение возросло с 61 565 до 182 814. Следует предположить, что по результатам индексации статей за второй квартал число публикаций также возрастет и тенденция роста возобновится.

Число публикаций

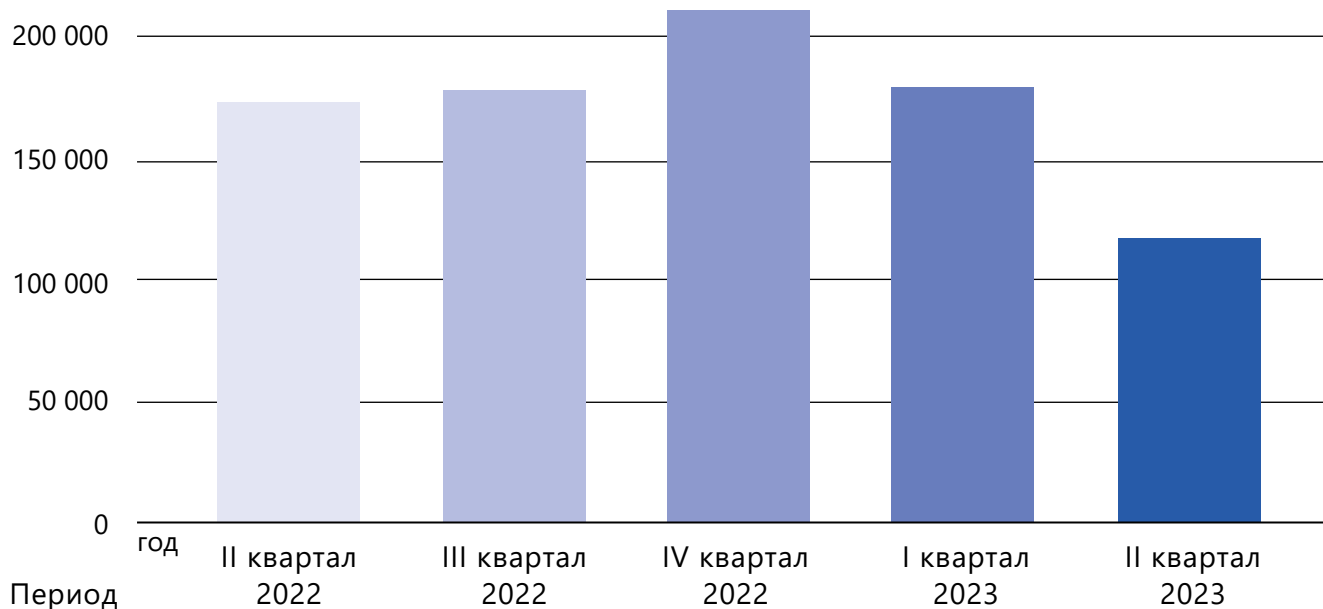


Рисунок 1. Динамика квартального числа публикаций за 2023 год

Число публикаций

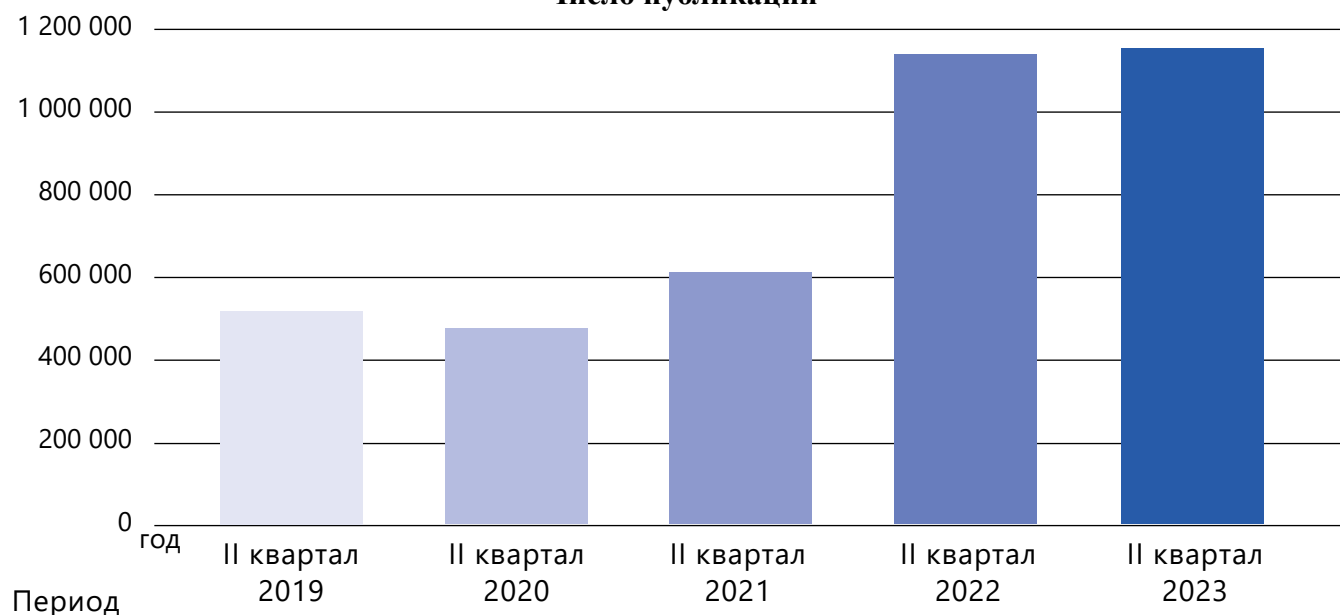


Рисунок 2. Динамика числа научных публикаций в сфере искусственного интеллекта в первом квартале за прошедшие пять лет

Ниже приведен рейтинг стран по числу научных публикаций в сфере искусственного интеллекта за второй квартал 2023 года:

- | | |
|--------------------|------------------------|
| 1. CHINA. | 11. NETHERLANDS. |
| 2. UNITED STATES. | 12. REPUBLIC OF KOREA. |
| 3. UNITED KINGDOM. | 13. INDONESIA. |
| 4. INDIA. | 14. SPAIN. |
| 5. GERMANY. | 15. SWITZERLAND. |
| 6. CANADA. | 16. RUSSIA. |
| 7. AUSTRALIA. | 17. SWEDEN. |
| 8. ITALY. | 18. BRAZIL. |
| 9. FRANCE. | 19. IRAN. |
| 10. JAPAN. | 20. SINGAPORE. |

Россия занимает 16-е место с 782 публикациями, опережая Швецию (685 статей), но уступая Швейцарии (830 статей). Динамика позиций стран-лидеров и России за последние пять кварталов приведена на рисунке 3.

**РОССИЯ ЗАНИМАЕТ
16-Е МЕСТО
С 782 ПУБЛИКАЦИЯМИ,
ОПЕРЕЖАЯ ШВЕЦИЮ
(685 СТАТЕЙ), НО УСТУПАЯ
ШВЕЙЦАРИИ (830 СТАТЕЙ).**

Таким образом, лидером по научно-публикационной активности в сфере искусственного интеллекта за второй квартал 2023 года является Китай, сохранявший свое место в рейтинге в течение последних четырех кварталов. При этом Россия находится на 16-м месте, демонстрируя укрепление своей позиции по отношению к прошлому кварталу, в котором Россия занимала 19-е место.

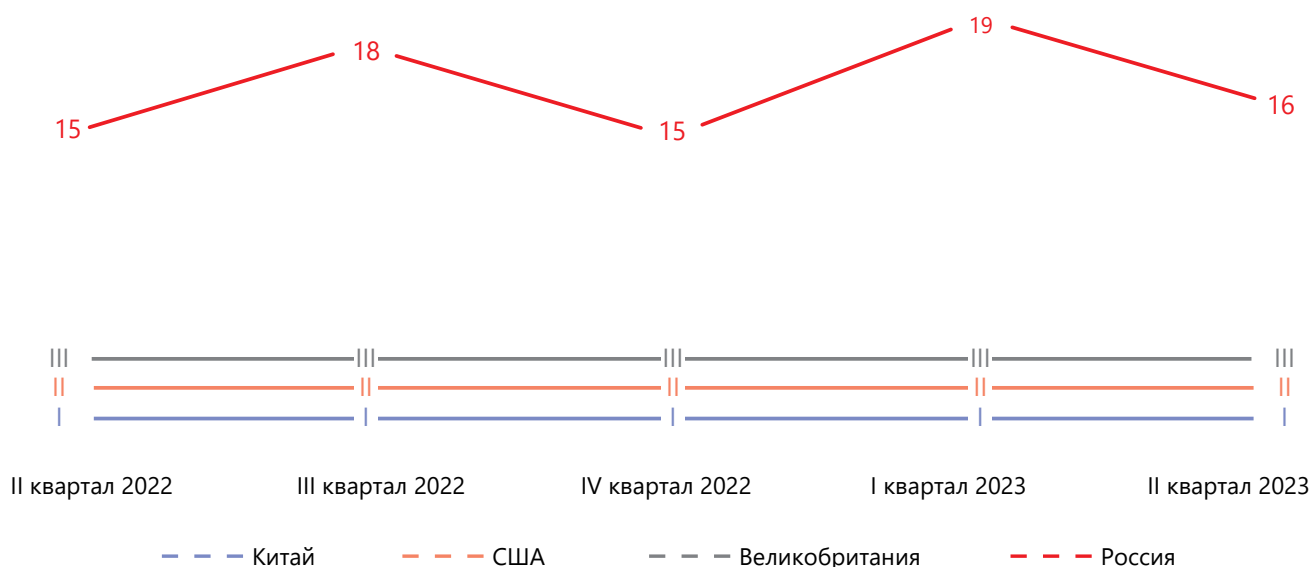


Рисунок 3. Позиции стран - лидеров по числу научных публикаций в сфере искусственного интеллекта в течение последних четырех кварталов

Темы публикаций

Наиболее популярными по числу публикаций за второй квартал 2023 года являются следующие темы:

1. General Medicine – 7219 публикаций.
2. Computer Science Applications – 6834 публикации.
3. Electrical and Electronic Engineering – 6173 публикации.
4. Software – 5069 публикаций.

При этом по среднему цитированию наиболее востребованы следующие темы:

1. Software.
2. Computer Science Applications.
3. Electrical and Electronic Engineering.
4. General Medicine.

Наиболее популярная тема по числу публикаций во втором квартале 2023 года – общая медицина. При этом наиболее востребованная тема по числу цитирований – программное обеспечение. В целом перечни из пяти самых популярных и самых востребованных тем соответствуют друг другу.

В рамках наиболее популярных тем научных публикаций в сфере искусственного интеллекта лидерами по числу публикаций являются страны, представленные в таблице 1.

Лидерами по числу научных публикаций в наиболее востребованных темах являются страны, приведенные в таблице 2.

Анализ тематики научных публикаций, вышедших во втором квартале 2023 года, показывает, что по числу публикаций в наиболее популярных темах лидируют Китай (наиболее активная организация – Китайская академия наук) и США (наиболее активные организации – университеты Нортирстен, Гарвард и Эй-энд-эм). Эти же страны и организации сохраняют свое лидерство в наиболее востребованных темах. Россия занимает самые высокие позиции в области общей медицины и электроники и электротехники.

Таблица 1.

Страны – лидеры по числу публикаций
и самые активные организации в этих странах
по наиболее популярным темам

Позиция	Страна	Название организации
Общая медицина		
1	США	Гарвардский университет
2	Китай	Китайская академия наук
3	Великобритания	Университетский колледж Лондона
4	Индонезия	Семарангский государственный университет
5	Канада	Университет Торонто
21	Россия	Московский государственный университет имени М. В. Ломоносова
Прикладная информатика		
1	Китай	Китайская академия наук
2	США	Мичиганский университет
3	Великобритания	Университетский колледж Лондона
4	Индия	Университет REVA
5	Канада	Университет Торонто
37	Россия	Московский государственный университет имени М. В. Ломоносова
Электротехника и электроника		
1	Китай	Китайская академия наук
2	США	Техасский университет Эй-энд-эм
3	Индия	Технологический институт Веллур
4	Великобритания	Имперский колледж Лондона
5	Республика Корея	Университет Корё
14	Россия	Технологический университет
Программное обеспечение		
1	Китай	Китайская академия наук
2	США	Университет Нортистерн
3	Индия	Университет SRM
4	Австралия	Университет Сиднея
5	Великобритания	Университет Оксфорд
31	Россия	Московский государственный университет имени М. В. Ломоносова

Таблица 2.

Страны – лидеры по числу публикаций
и самые активные организации в этих странах
по наиболее востребованным темам

Позиция	Страна	Название организации
Программное обеспечение		
1	Китай	Китайская академия наук
2	США	Университет Нортистерн
3	Индия	Университет SRM
4	Австралия	Университет Сиднея
5	Великобритания	Университет Оксфорд
31	Россия	Московский государственный университет имени М. В. Ломоносова
Прикладная информатика		
1	Китай	Китайская академия наук
2	США	Мичиганский университет
3	Великобритания	Университетский колледж Лондона
4	Индия	Университет REVA
5	Канада	Университет Торонто
37	Россия	Московский государственный университет имени М. В. Ломоносова
Электротехника и электроника		
1	Китай	Китайская академия наук
2	США	Техасский университет Эй-энд-эм
3	Индия	Технологический институт Веллур
4	Великобритания	Имперский колледж Лондона
5	Республика Корея	Университет Корё
14	Россия	Технологический университет
Общая медицина		
1	США	Гарвардский университет
2	Китай	Китайская академия наук
3	Великобритания	Университетский колледж Лондона
4	Индонезия	Семарангский государственный университет
5	Канада	Университет Торонто
21	Россия	Московский государственный университет имени М. В. Ломоносова

Наиболее значимые научные результаты

Наиболее значимые результаты, как правило, публикуются исследователями в научных изданиях, характеризующихся высоким уровнем признания мировым сообществом. Уровень признания научного издания обычно определяется средней цитируемостью опубликованных в нем статей, количественно характеризующейся квартилем цитируемости и импакт-фактором журнала. Данные метрики публикуются на сайте Scimago Journal Ranking.

Анализ библиографических данных научных статей в сфере искусственного интеллекта, опубликованных за второй квартал 2023 года, показывает, что доля статей в журналах первого квартиля цитируемости (Q1) от общего числа публикаций составляет 0,7%.

На основе указанных выше метрик сформирован перечень из 9 статей, опубликованных за рассматриваемый период в наиболее авторитетных научных журналах:

1. Iuliu Alexandru Zamfirache, Radu-Emil Precup, Raul-Cristian Roman, Emil M Petriu. Neural Network-based control using Actor-Critic Reinforcement Learning and Grey Wolf Optimizer with experimental servo system validation // Expert Systems with Applications, Volume: 225, Pages: 120112–120112 DOI: 10.1016/j.eswa.2023.120112 Число цитирований: 2. В этой статье представлен новый подход к управлению, сочетающий в себе обучение с подкреплением актера и критики (RL), и оптимизатор «серых волков»¹ (GWO). Алгоритм GWO заменяет алгоритм градиентного спуска (GD) для оптимизации Critic, устраняя его медленную сходимость и проблемы с локальным оптимумом.
2. Rachel L Moran, Emilie J Richards, Claudia Patricia Ornelas-García, Joshua B Gross, Alexandra Donny, Jonathan Wiese, Alex C Keene, Johanna E Kowalko, Nicolas Rohner, Suzanne E McGaugh // Selection-driven trait loss in independently evolved cavefish populations // Nature communications, Volume: 14, Issue: 1, Pages: 2557. DOI: 10.1038/s41467-023-37909-8 Число цитирований: 2. Авторы используют повторное секвенирование всего генома мексиканской тетры для изучения повторной адаптации у пещерных рыб. Отбор по генетической изменчивости и мутации de novo способствует изменению признаков. Подход использует собственную декомпозицию для сканирования генома на предмет изменений частоты аллелей.
3. Cheyenne Ziegler, Jonathan Martin, Claude Sinner, Faruck Morcos // Latent generative landscapes as maps of functional diversity in protein sequence space //

УРОВЕНЬ ПРИЗНАНИЯ НАУЧНОГО ИЗДАНИЯ ОБЫЧНО ОПРЕДЕЛЯЕТСЯ СРЕДНЕЙ ЦИТИРУЕМОСТЬЮ ОПУБЛИКОВАННЫХ В НЕМ СТАТЕЙ, КОЛИЧЕСТВЕННО ХАРАКТЕРИЗУЮЩЕЙСЯ КВАРТИЛЕМ ЦИТИРУЕМОСТИ И ИМПАКТ-ФАКТОРОМ ЖУРНАЛА.

Nature communications, Volume: 14, Pages: 2222 DOI: 10.1038/s41467-023-37958-z. Число цитирований: 2.

2. Вариационные автоэнкодеры классифицируют последовательности белков и генерируют новые, сохраняя при этом статистические свойства. Это исследование сосредоточено на скрытом многообразии, встроенном в информацию о последовательности, с использованием прямого анализа связи и гамильтоновой модели Поттса.
4. Jeffrey A Ruffolo 1, 2, Lee-Shin Chu 3, 2, Sai Pooja Mahajan 3, 2, Jeffrey J Gray // Fast, accurate antibody structure prediction from deep learning on massive set of natural antibodies // Nature communications, Volume: 14, Pages: 2389 DOI: 10.1038/s41467-023-38063-x. Число цитирований: 2. Авторы представляют быстрый метод глубокого обучения IgFold для прогнозирования структуры антител. Используя предварительно обученную языковую модель и графовые сети, IgFold предсказывает координаты атомов основной цепи с высокой точностью менее чем за 25 секунд.
5. Q Xiong 1, 2, M R Brudzinski 3, D Gossett 3, Q Lin 4, 5, J C Hampton // Seismic magnitude clustering is prevalent in field and laboratory catalogs // Nature communications, Volume: 14, Pages: 2056 DOI: 10.1038/s41467-023-37782-5 Число цитирований: 2. Это исследование выявляет статистически значимую кластеризацию величин в различных полевых и лабораторных каталогах в разных пространственных масштабах.
6. Wenjing Ma, Jiaying Lu, Hao Wu // Cellcano: supervised cell type identification for single cell ATAC-seq data // Nature communications, Volume: 14, Pages: 1864 DOI: 10.1038/s41467-023-37439-3 Число цитирований: 2. Авторы разрабатывают вычислительный метод Cellcano, основанный на двухэтапном алгоритме контролируемого обучения для определения типов клеток из данных scATAC-seq. Метод смягчает сдвиг распределения между эталонными и целевыми данными и повышает эффективность прогнозирования.
7. Xiaoxun Gong, He Li, Nianlong Zou, Runzhang Xu, Wenhui Duan, Yong Xu // General framework for E (3) -equivariant neural network representation of density functional theory Hamiltonian // Nature communications, Volume: 14, Pages: 2848 DOI: 10.1038/s41467-023-38468-8 Число цитирований: 2. Авторы предлагают E (3) -эквивариантную структуру глубокого обучения для представления гамильтониана теории функционала плотности (DFT) как функции структуры материала,

¹ Алгоритм оптимизации «серых волков» – это один из недавних биоинспирированных алгоритмов оптимизации, основанный на имитации загонной охоты стаи серых волков.

ЛИДЕРОМ ПО ПУБЛИКАЦИЯМ В НАИБОЛЕЕ АВТОРИТЕТНЫХ ЖУРНАЛАХ ЗА ВТОРОЙ КВАРТАЛ 2023 ГОДА СТАЛ КИТАЙ.

которая может естественным образом сохранять евклидову симметрию даже при наличии спин-орбитальной связи.

8. Firas Al-Hindawi, Tejaswi Soori, Han Hu, Md Mahfuzur Rahman Siddiquee, Hyunsoo Yoon, Teresa Wu, Ying Sun // A framework for generalizing critical heat flux detection models using unsupervised image-to-image translation // Expert Systems with Applications, Volume: 227, Pages: 120265–120265 DOI: 10.1016/j.eswa.2023.120265 Число цитирований: 2. Авторы предлагают новую неконтролируемую структуру, которая повышает обобщаемость и адаптируемость обученных моделей обнаружения CHF. Данный подход использует неконтролируемую модель преобразования изображения в изображение (U2I) для преобразования изображений целевого набора данных, чтобы они напоминали домен, на котором модель была первоначально обучена.
9. Zhonghua Liu, Zhihui Lai, Weihua Ou, Kaibing Zhang, Hua Huo // Discriminative sparse least square regression for semi-supervised learning // Information Sciences, Volume: 636, Pages: 118903–118903 DOI: 10.1016/j.ins.2023.03.128 Число цитирований: 2. Авторы используют оценочную метку наблюдаемых выборок для разработки новой целевой функции в форме обобщенной регрессии, дополнительно обобщая предыдущую структуру

регрессии наименьших квадратов. Новый совместно разреженный регуляризованный термин предназначен для полного использования оценочной информации о метках, что, как ожидается, заставит извлеченные признаки каждого класса быть совместно разреженными вместо того, чтобы изученная проекция была совместно разреженной.

Параметры научных изданий, опубликовавших данные статьи, приведены в таблице 3.

Таблица 3. Пять наиболее авторитетных журналов, в которых опубликованы статьи в сфере искусственного интеллекта, и тематика этих журналов

Название журнала	Импакт-фактор	Квартиль
Information Sciences	2,29	1
Nature Communications	2,25	1
Future Generation Computer Systems	1,88	1
Expert Systems with Applications	1,87	1
Computers in Biology and Medicine	1,22	1

Из таблицы 3 следует, что наиболее авторитетным научным журналом, опубликовавшим статьи в сфере искусственного интеллекта за второй квартал 2023 года, является Information Sciences, отнесенный к квартилью цитируемости 1 и характеризующийся импакт-фактором 2,29 (SJR). Основная тематика журнала – информатика, изучение искусственного интеллекта и технологий поддержки принятия решений.

Лидером по публикациям в наиболее авторитетных журналах за второй квартал 2023 года стал Китай.

ИНТЕРЕСЫ ИНДУСТРИИ

Отдельный интерес представляют научные исследования, в которых принимают участие коммерческие компании, поскольку результаты данных исследований с высокой долей вероятности будут внедрены в компаниях и выйдут на рынок в качестве продукта. Анализ научных публикаций, подготовленных в соавторстве с сотрудниками компаний, показал, что за рассматриваемый период интерес индустрии привлекали следующие темы:

1. Computer Science Applications – 52 публикации.
2. Software – 36 публикаций.
3. General Physics and Astronomy – 33 публикации.
4. Education – 29 публикаций.
5. Industrial and Manufacturing Engineering – 28 публикаций.
6. General Medicine – 21 публикация.
7. History – 21 публикация.
8. Computer Networks and Communications – 19 публикаций.
9. Applied Mathematics – 18 публикаций.

Ниже приведен перечень компаний – лидеров по числу публикаций в сфере искусственного интеллекта.

1. Google – 130 публикаций.
2. Virginia Tech – 124 публикации.
3. IBM – 106 публикаций.
4. Shanghai Electric – 62 публикации.
5. Riken – 58 публикаций.
6. State Grid Corporation of China – 58 публикаций.
7. China Southern Power Grid – 43 публикации.
8. AstraZeneca – 33 публикации.
9. China General Nuclear Power Corporation – 26 публикаций.
10. Alibaba Group – 18 публикаций.

Таким образом, коммерческие компании публикуют результаты научных исследований по следующим темам: прикладная информатика, программное обеспечение, общая физика и математика, искусственный интеллект, общая медицина, компьютерные сети. Наиболее активной организацией является Google со 130 публикациями, число которых по отношению к первому кварталу 2023 года (60) увеличилось более чем в два раза. Данные результаты опубликованы в журналах, не входящих в перечень наиболее авторитетных.

ПРИЛОЖЕНИЕ

МЕТОДОЛОГИЯ

1. Поиск и анализ научных публикаций в сфере искусственного интеллекта проводился на интернет-ресурсе Lens.org.
2. Поиск публикаций проводился 29 июня 2023 г. по дате публикации из интервала 1 апреля 2023 г. – 29 июня 2023 г.
3. Поиск публикаций проводился по следующим областям знаний:
 - artificial intelligence;
 - machine learning;
 - deep learning;
 - computer vision;
 - pattern recognition;
 - artificial neural network;
 - convolutional neural network;
 - natural language processing;
 - reinforcement learning;
 - data science;
 - robustness (computer science);
 - speech recognition. ✨