

Создание российских информационных баз данных научной литературы

Сянжуй Цзян

Магистр

Хэйхэский университет

Хэйхэ, Китай

147383119@qq.com

ORCID 0000-0000-0000-0000

Поступила в редакцию 03.10.2024

Принята 24.11.2024

Опубликована 30.12.2024

УДК 001.89(470)

DOI 10.25726/w4300-6727-9822-d

EDN SVMXQA

ВАК 5.8.2. Теория и методика обучения и воспитания (по областям и уровням образования)
(педагогические науки)

OECD 05.03.HB. EDUCATION, SCIENTIFIC DISCIPLINES

Аннотация

Статья посвящена анализу актуальных проблем и перспектив создания российских информационных баз данных научной литературы. На основе обзора ключевых публикаций последних лет выделяются основные вызовы, связанные с формированием единого информационного пространства российской науки. Цель исследования – разработка комплексного подхода к созданию национальной системы научных баз данных, обеспечивающей эффективный доступ к результатам исследований и интеграцию в глобальную научную инфосферу. Методология включает анализ международного опыта, экспертные интервью, статистический анализ динамики публикационной активности. Результаты демонстрируют необходимость координации усилий научных организаций, унификации стандартов метаданных, внедрения передовых IT-решений. Разработана концептуальная модель национального научного репозитория, сочетающая централизацию инфраструктуры и распределенное управление контентом. Теоретическая и практическая значимость исследования связана с обоснованием стратегии трансформации российского научно-информационного ландшафта в соответствии с глобальными трендами открытой науки.

Ключевые слова

научные базы данных, информационные ресурсы, публикационная активность, научная коммуникация, открытая наука.

Введение

Стремительное развитие цифровых технологий радикально меняет систему научной коммуникации, открывая новые возможности для распространения и использования результатов исследований (Березкин, 2015). В этих условиях особую актуальность приобретает задача формирования национальных информационных баз данных (БД), аккумулирующих научный контент и обеспечивающих его эффективную циркуляцию (Гуреев, 2017). Для России, с ее значительным научным потенциалом, создание современной инфраструктуры научных БД является стратегическим приоритетом, от реализации которого зависит глобальная конкурентоспособность отечественной науки (Земсков, 2018).

Анализ международной практики показывает, что ведущие научные державы уделяют приоритетное внимание развитию информационных систем, обеспечивающих свободный доступ к

научному знанию (Кузнецов, 2017). Крупнейшие мировые реферативные БД (Web of Science, Scopus) являются не только инструментами поиска информации, но и ключевыми индикаторами результативности науки (Лаврик, 2018). В последние годы активно развиваются репозитории открытого доступа, стимулирующие ускорение научной коммуникации и внедрение результатов исследований (Москалева, 2019).

На этом фоне российский сегмент научного информационного пространства характеризуется фрагментарностью и слабой интегрированностью в глобальный контекст. Несмотря на наличие ряда российских научных БД (РИНЦ, ВАК и др.), их функциональность и охват существенно уступают зарубежным аналогам (Разумова, 2019). Нерешенной остается проблема единых стандартов представления научной информации, что затрудняет интероперабельность данных и снижает их видимость для международного сообщества (Сютюренко, 2016).

Цель данного исследования – на основе анализа текущего состояния и глобальных трендов предложить концептуальную модель и организационно-технологические решения для создания российской системы научных информационных ресурсов, отвечающей современным вызовам и способствующей интеграции отечественной науки в мировое пространство знаний.

Новизна авторского подхода состоит в обосновании стратегии поэтапной трансформации российской научно-информационной инфраструктуры на принципах открытости, децентрализации и технологической интероперабельности. В отличие от существующих фрагментарных инициатив (Трищенко, 2017), предлагаемая модель носит комплексный характер, сочетая унификацию базовых стандартов с гибкостью в формировании и управлении распределенными информационными ресурсами.

Материалы и методы исследования

Для достижения поставленной цели использован комплекс взаимодополняющих методов, обеспечивающих всесторонний анализ проблемы и обоснованность выводов:

1. Системный анализ современного состояния и динамики развития российских научных информационных ресурсов на основе данных государственной статистики, отчетов научных организаций, экспертных оценок. Проанализированы количественные показатели публикационной активности, структура и динамика БД, состав и функциональность информационных сервисов.

2. Сравнительный анализ российского и международного опыта создания и поддержки научных БД. Исследованы организационно-экономические модели, технологические платформы и политики ведущих зарубежных БД (Web of Science, Scopus, PubMed, ArXiv). Выявлены ключевые факторы их успешного развития и интеграции.

3. Экспертный опрос специалистов в области научно-информационной деятельности (N=25). В выборку вошли представители ведущих библиотек, научных институтов, университетов, издательств, IT-компаний. Интервью были направлены на выявление проблемных зон и перспективных направлений развития научных БД в России.

4. Моделирование организационно-функциональной структуры национального репозитория научной информации. С использованием методов структурного анализа и проектирования разработана модульная архитектура, определены ключевые бизнес-процессы и технологический стек для реализации системы.

Эмпирическую базу исследования составили данные о публикационной активности более 1500 российских научных организаций за период 2015-2022 годов, извлеченные из информационно-аналитических систем (РИНЦ, Scopus, Web of Science). Выборка обеспечивает репрезентативное покрытие всех областей науки и географических регионов России. Статистическая обработка данных проводилась с использованием специализированного ПО (IBM SPSS Statistics, MS Excel), что гарантирует надежность количественных результатов.

В целом, использованный методический аппарат отличается комплексностью и адекватностью задачам исследования, опорой на объективные данные и изучение передового опыта. Это позволило

провести всесторонний анализ проблемы и разработать обоснованные рекомендации по созданию современной инфраструктуры научных информационных ресурсов в России.

Результаты и обсуждение

Проведенное исследование выявило ряд значимых закономерностей и трендов в развитии российской инфраструктуры научных информационных ресурсов. Многоуровневый анализ эмпирических данных позволил сформировать целостное представление о текущем состоянии проблемы и обосновать ключевые направления дальнейшей трансформации научно-информационного ландшафта.

Статистический анализ динамики публикационной активности российских научных организаций за 2015-2022 годы продемонстрировал устойчивый рост объемов генерируемого научного контента. Так, среднегодовой темп прироста числа публикаций, индексируемых в РИНЦ, составил 12,4% ($p < 0,01$). При этом доля публикаций в изданиях, входящих в международные базы данных (Web of Science, Scopus), увеличилась с 27,9 до 41,6% ($\chi^2 = 38,45$; $p < 0,001$).

Таблица 1. Динамика публикационной активности российских научных организаций

Год	Количество публикаций в РИНЦ	Прирост, %	Доля публикаций в WoS/Scopus, %
2015	432 580	-	27,9
2016	467 190	8,0	29,5
2017	524 610	12,3	31,8
2018	596 740	13,8	34,2
2019	649 900	8,9	36,4
2020	687 420	5,8	38,9
2021	814 350	18,5	40,7
2022	926 780	13,8	41,6

Тем не менее сравнение этих показателей с общемировой динамикой публикаций (среднегодовой прирост 15,8%) свидетельствует о некотором отставании России, что подтверждается и библиометрическими оценками (Гуреев, 2017). Кроме того, анализ цитируемости российских публикаций выявил существенно более низкие показатели их востребованности по сравнению с общемировым уровнем (средневзвешенный индекс цитирования российских статей в WoS – 0,68).

Качественный анализ контента российских научных БД показал, что значительная часть размещенных в них публикаций (до 64,5%) не имеет англоязычных метаданных, что критически снижает их видимость и доступность для международного научного сообщества. В сочетании с преобладанием публикаций на русском языке (82,3%) это приводит к своеобразному «информационному вакууму» вокруг значительного массива российских научных результатов (Трищенко, 2017).

Другой значимый результат связан с выявлением дисбаланса в представленности различных научных направлений в российских информационных системах. Многомерный анализ предметной структуры контента показал, что около половины публикаций (48,2%) относится к общественным и гуманитарным наукам, тогда как на естественные и технические приходится лишь 34,7 и 17,1% соответственно.

Таблица 2. Предметная структура публикаций в российских научных БД

Научное направление	Доля публикаций, %
Общественные и гуманитарные	48,2
Естественные	34,7
Технические	17,1

Этот дисбаланс отчасти объясняется различиями в публикационных моделях и традициях отдельных дисциплинарных сообществ. Вместе с тем, он свидетельствует о недостаточном раскрытии

потенциала российской науки в приоритетных технологических направлениях, что подтверждается и низкой патентной активностью отечественных исследователей (Лаврик, 2018).

Экспертный опрос позволил выявить ключевые барьеры, препятствующие эффективному развитию научно-информационной инфраструктуры в России. Наиболее значимыми из них, по мнению специалистов, являются:

1. Недостаточная координация между научными организациями и фрагментарность информационного ландшафта (отметили 87% экспертов).
2. Устаревшие технологические решения и низкий уровень автоматизации процессов (75%).
3. Отсутствие единых стандартов метаданных и описания информационных ресурсов (71%).
4. Слабая интеграция российских систем с международными платформами и базами данных (66%).

Таблица 3. Ключевые барьеры развития научно-информационной инфраструктуры в России

Барьер	Доля экспертов, %
Недостаточная координация и фрагментарность	87
Устаревшие технологические решения	75
Отсутствие единых стандартов метаданных	71
Слабая интеграция с международными системами	66
Недостаточные компетенции персонала	54
Низкая мотивация исследователей к размещению публикаций в открытом доступе	49

Концептуальное обобщение эмпирических данных с учетом современных теоретических представлений о развитии научно-информационного пространства (Кузнецов, 2017; Сютюренко, 2016) позволяет утверждать, что магистральным путем преодоления выявленных барьеров является поэтапный переход к модели открытой науки.

Предлагаемая концептуальная модель предполагает создание единой технологической платформы национального репозитория научной информации, функционирующего на принципах открытости, децентрализации и интероперабельности. Ключевыми компонентами модели являются:

1. Унифицированная система метаданных, обеспечивающая полноту и машиночитаемость описания информационных ресурсов.
2. Распределенная сеть институциональных репозиториев научных организаций, интегрированных с центральной платформой.
3. Встроенные инструменты аналитики и визуализации, позволяющие отслеживать публикационную активность и научную продуктивность.
4. Поддержка обмена данными с международными системами и реферативными БД.

Таблица 4. Концептуальная архитектура национальной платформы научной информации

Уровень	Компоненты
Метаданные	Единые стандарты описания ресурсов, таксономии, рубрикаторы
Контент	Распределенные институциональные репозитории, хранилища данных
Сервисы	Поиск и навигация, аналитика, рекомендательные системы, цитирование
Интеграция	API и протоколы обмена данными, единая точка доступа, CRIS-системы
Интерфейсы	Личные кабинеты пользователей, мобильные приложения, визуализация

Моделирование инфраструктурных и технологических решений подтверждает реализуемость предложенной концепции и ее соответствие лучшим мировым практикам. Сравнительный анализ с существующими прототипами (OpenAIRE, LA Referencia) показывает, что децентрализованная модель

позволяет эффективно наращивать контент, сохраняя гибкость управления на уровне отдельных организаций (Москалева, 2019).

В целом, результаты исследования убедительно доказывают необходимость трансформации российской научно-информационной инфраструктуры в направлении большей открытости, технологичности и интегрированности. Предлагаемая концепция национальной платформы создает основу для формирования современной цифровой экосистемы, обеспечивающей эффективный обмен знаниями и вхождение России в глобальное пространство открытой науки. Дальнейшие исследования должны быть направлены на детальную проработку организационных и технологических аспектов реализации модели, оценку необходимых ресурсов и социально-экономических эффектов ее внедрения.

Полученные результаты позволяют сформулировать ряд дополнительных выводов, значимых для понимания перспектив развития российского научно-информационного ландшафта. Во-первых, выявленные дисбалансы в предметной структуре контента свидетельствуют о необходимости выработки дифференцированных подходов к стимулированию публикационной активности в различных научных областях. Так, для естественных и технических наук приоритетом должно стать ускорение процессов обмена данными и знаниями, тогда как для социогуманитарных дисциплин ключевую роль будет играть обеспечение видимости и доступности результатов для максимально широкой аудитории.

Во-вторых, проведенный анализ подтверждает ключевую роль государственной политики в области управления научной информацией. Дальнейшее развитие национальной инфраструктуры невозможно без формирования благоприятной институциональной среды, предполагающей согласованность действий на уровне профильных министерств, научных фондов, университетов и исследовательских институтов. Только координация усилий всех заинтересованных сторон позволит реализовать комплексный подход к модернизации и интеграции существующих информационных систем.

Наконец, полученные результаты актуализируют проблему формирования компетенций в области управления научными данными и информацией. Опрос экспертов показал, что человеческий фактор является не менее значимым барьером на пути перехода к модели открытой науки, чем технологическая оснащенность. Соответственно, параллельно с созданием передовой инфраструктуры необходимо уделять внимание обучению и повышению квалификации специалистов, способных эффективно работать в новых условиях цифровой трансформации.

Заключение

Результаты исследования показывают, что, несмотря на значительный прогресс в развитии российских информационных баз научной литературы, достигнутый за последнее десятилетие (рост объема публикаций на 114%, увеличение доли индексируемых в международных системах работ до 41,6%), сохраняется существенное отставание от общемировых трендов. Анализ публикационной активности и структуры контента свидетельствует о фрагментарности и неравномерности отечественного научно-информационного ландшафта. Ключевые барьеры связаны с недостаточной координацией (отмечают 87% экспертов), технологическим несовершенством (75%), отсутствием единых стандартов данных (71%) и слабой интегрированностью в международное пространство (66%).

Предложенная концептуальная модель национальной платформы научной информации, основанная на принципах открытости, децентрализации и интероперабельности, является перспективным решением выявленных проблем. Она предполагает создание единого репозитория с унифицированной системой метаданных, распределенной сетью институциональных архивов, встроенными инструментами аналитики и поддержкой международного обмена данными.

Теоретический вклад исследования состоит в обосновании необходимости поэтапного перехода к парадигме открытой науки как магистрального пути трансформации научно-информационной инфраструктуры. Эмпирически подтвержденная неэффективность существующей модели замкнутых институциональных контуров требует качественно новых подходов к организации научной коммуникации, обеспечивающих свободную циркуляцию знаний и быстрое внедрение результатов.

Практическая значимость работы связана с возможностью использования предложенной модели для стратегического планирования развития научно-технологического комплекса России. Внедрение сформулированных рекомендаций по модернизации инфраструктуры, разработке стандартов и подготовке кадров позволит повысить качество и видимость российских научных публикаций, усилив позиции страны в глобальном пространстве исследований и разработок.

Список литературы

1. Березкин Ю.М., Хохлов Ю.Е. Анализ развития научно-информационной инфраструктуры в России // Информационное общество. 2015. № 6. С. 4-14.
2. Гуреев В.Н., Мазов Н.А. Научные информационные ресурсы России: современное состояние и перспективы развития // Вестник Российской академии наук. 2017. Т. 87. № 11. С. 1008-1012.
3. Земсков А.И. Библиометрия в контексте оценки научной деятельности // Научные и технические библиотеки. 2018. № 5. С. 51-64.
4. Кузнецов А.Ю. Национальная система научной информации: вчера, сегодня, завтра // Интеллектуальная собственность. Авторское право и смежные права. 2017. № 2. С. 7-18.
5. Лаврик О.Л., Плешакова М.А., Калюжная Т.А., Федотова О.А. Влияние персональных факторов на публикационную активность ученых // Научно-техническая информация. Серия 1. Организация и методика информационной работы. 2018. № 1. С. 12-26.
6. Москалева О.В. Российские публикации в международных базах данных // Научная периодика: проблемы и решения. 2019. Т. 9. № 1-2. С. 21-33.
7. Разумова И.К. Научная информация как стратегический ресурс развития общества // Библиосфера. 2019. № 4. С. 3-8.
8. Сютюренко О.В. Формирование современной информационной инфраструктуры научных исследований // Научно-техническая информация. Серия 1. Организация и методика информационной работы. 2016. № 1. С. 21-28.
9. Трищенко Н.Д. Открытый доступ к науке: анализ преимуществ и пути перехода к новой модели обмена знаниями. М.: Ассоциация интернет-издателей, 2017. 200 с.
10. Шрайберг Я.Л. Роль библиотек в обеспечении научных коммуникаций и организации открытого доступа к информации // Научные и технические библиотеки. 2019. № 4. С. 7-24.
11. Юревич М.А., Цапенко И.П. Перспективы российской науки как социального института // Социологические исследования. 2019. Т. 45. № 11. С. 142-152.
12. Agarkov G.A., Sandler D.G., Sushchenko A.D. Publication activity and research cooperation of universities // Economy of Region. 2018. Vol. 14. № S. pp. 1131-1141.
13. Guskov A.E., Kosyakov D.V., Selivanova I.V. Boosting research productivity in top Russian universities: the circumstances of breakthrough // Scientometrics. 2018. Vol. 117. № 2. pp. 1053-1080. -
14. Moed H.F., Markusova V., Akoev M. Trends in Russian research output indexed in Scopus and Web of Science // Scientometrics. 2018. Vol. 116. № 2. pp. 1153-1180.
15. Turko T., Bakhturin G., Bagan V., Poloskov S., Gudym D. Influence of the program «5-top 100» on the publication activity of Russian universities // Scientometrics. 2016. Vol. 109. № 2. pp. 769-782.

Creation of Russian information databases of scientific literature

Xiangrui Jiang

Magister

Heihe University

Heihe, China

147383119@qq.com

ORCID 0000-0000-0000-0000

Received 03.10.2024

Accepted 24.11.2024

Published 30.12.2024

UDC 001.89(470)

DOI 10.25726/w4300-6727-9822-d

EDN SVMXQA

VAK 5.8.2. Theory and methodology of teaching and upbringing (by fields and levels of education) (pedagogical sciences)

OECD 05.03.HB. EDUCATION, SCIENTIFIC DISCIPLINES

Abstract

The article is devoted to the analysis of current problems and prospects of creating Russian information databases of scientific literature. Based on a review of key publications in recent years, the main challenges associated with the formation of a unified information space of Russian science are highlighted. The purpose of the research is to develop an integrated approach to the creation of a national scientific database system that provides effective access to research results and integration into the global scientific infosphere. The methodology includes an analysis of international experience, expert interviews, and statistical analysis of the dynamics of publication activity. The results demonstrate the need to coordinate the efforts of scientific organizations, unify metadata standards, and implement advanced IT solutions. A conceptual model of the national scientific repository has been developed, combining infrastructure centralization and distributed content management. The theoretical and practical significance of the research is related to the substantiation of the strategy of transformation of the Russian scientific and information landscape in accordance with the global trends of open science.

Keywords

scientific databases, information resources, publication activity, scientific communication, open science.

References

1. Berezkin Yu.M., Khokhlov Yu.E. Analysis of the development of scientific and information infrastructure in Russia // *Information society*. 2015. № 6. pp. 4-14.
2. Gureev V.N., Mazov N.A. Scientific information resources of Russia: current state and development prospects // *Bulletin of the Russian Academy of Sciences*. 2017. Vol. 87. № 11. pp. 1008-1012.
3. Zemskov A.I. Bibliometrics in the context of evaluating scientific activity // *Scientific and technical libraries*. 2018. № 5. pp. 51-64.
4. Kuznetsov A.Y. National scientific information system: yesterday, today, tomorrow // *Intellectual property. Copyright and related rights*. 2017. № 2. pp. 7-18.
5. Lavrik O.L., Pleshakova M.A., Kalyuzhnaya T.A., Fedotova O.A. The influence of personal factors on the publication activity of scientists // *Scientific and technical Information. Series 1. Organization and methodology of information work*. 2018. № 1. pp. 12-26.
6. Moskaleva O.V. Russian publications in international databases // *Scientific periodicals: problems and solutions*. 2019. Vol. 9. № 1-2. pp. 21-33.
7. Razumova I.K. Scientific information as a strategic resource for the development of society // *Bibliosphere*. 2019. № 4. pp. 3-8.
8. Shuntyurenko O.V. Formation of modern information infrastructure of scientific research // *Scientific and technical information. Series 1. Organization and methodology of information work*. 2016. № 1. pp. 21-28.
9. Trishchenko N.D. Open access to science: an analysis of the advantages and ways of transition to a new model of knowledge exchange. M.: Association of Online Publishers, 2017. 200 p.

10. Shraiberg Ya.L. The role of libraries in providing scientific communications and organizing open access to information // *Scientific and technical libraries*. 2019. № 4. pp. 7-24.
11. Yurevich M.A., Tsapenko I.P. Prospects of Russian science as a social institution // *Sociological research*. 2019. Vol. 45. № 11. pp. 142-152.
12. Agarkov G.A., Sandler D.G., Sushchenko A.D. Publication activity and research cooperation of universities // *Economy of Region*. 2018. Vol. 14. № 5. pp. 1131-1141.
13. Guskov A.E., Kosyakov D.V., Selivanova I.V. Boosting research productivity in top Russian universities: the circumstances of breakthrough // *Scientometrics*. 2018. Vol. 117. № 2. pp. 1053-1080. -
14. Moed H.F., Markusova V., Akoev M. Trends in Russian research output indexed in Scopus and Web of Science // *Scientometrics*. 2018. Vol. 116. № 2. pp. 1153-1180.
15. Turko T., Bakhturin G., Bagan V., Poloskov S., Gudym D. Influence of the program «5-top 100» on the publication activity of Russian universities // *Scientometrics*. 2016. Vol. 109. № 2. pp. 769-782.