

ЭЛЕКТРОННЫЕ РЕСУРСЫ ОТКРЫТОГО ДОСТУПА. ОТКРЫТЫЕ АРХИВЫ ИНФОРМАЦИИ

УДК 002.1 – 021.341 + 025.4.03

<https://doi.org/10.33186/1027-3689-2026-5-146-163>

Метаданные и схемы метаданных в научных ресурсах открытого доступа

Н. С. Редькина

*ГПНТБ СО РАН, Новосибирск, Российская Федерация,
redkina@spsl.nsc.ru, <https://orcid.org/0000-0002-3486-9711>*

Аннотация. Качественные, стандартизованные метаданные, описанные по строгой схеме, делают результаты научных исследований легко находимыми и понятными как для пользователей, так и для информационно-поисковых систем. Создание качественных метаданных, отражаемых в научных ресурсах открытого доступа (РОД), и их гармонизация с помощью схем и стандартов – важнейшая задача, решение которой способствует повышению видимости и цитируемости результатов исследований, эффективности поиска и развитию научных направлений, сотрудничеству, изучению тематических структур. Целью исследования стало определение спектра метаданных и схем метаданных, отраженных в РОД. Для достижения поставленной цели были решены следующие задачи: 1) рассмотрен перечень полей метаданных, схем и стандартов метаданных, 2) определены особенности отражения метаданных для публикаций, индексируемых в крупнейших научных РОД (Google Scholar, OpenAlex, Dimensions и Lens, а также на российских ресурсах РИНЦ и Cobalt). Сделан вывод о перспективах развития ресурсов для поиска и обработки научной информации и проведения исследования с учетом постоянно развивающегося функционала, большей видимости и более высоких показателей точности и полноты метаданных. Устранение неоднозначности в описании за счет использования идентификаторов и единых стандартов описания, связывание метаданных из разных ресурсов, ввод дополнительных полей метаданных, позволяющих оценивать результаты исследований и систематизировать в соответствии с принятой в РОД классификацией, обеспечат поиск релевантной информации и упростят управление документами в различных поисковых системах.

Исследование выполнено в рамках реализации научного проекта ГПНТБ СО РАН (2022–2026 гг.) «Разработка модели функционирования научной библиотеки в информационной экосистеме открытой науки» № 122041100150-3.

Ключевые слова: открытый доступ, ресурсы, метаданные, стандарты, схемы, библиометрические показатели, альтметрия, классификация, индексирование

Для цитирования: Редькина Н. С. Метаданные и схемы метаданных в научных ресурсах открытого доступа // Научные и технические библиотеки. 2026. № 5. С. 146–163. <https://doi.org/10.33186/1027-3689-2026-5-146-163>

OPEN ACCESS DIGITAL RESOURCES. OPEN INFORMATION ARCHIVES

UDC 002.1 – 021.341 + 025.4.03

<https://doi.org/10.33186/1027-3689-2026-5-146-163>

Metadata and metadata schemes in open access research resources

Natalya S. Redkina

*SB RAS State Public Scientific and Technological Library,
Novosibirsk, Russian Federation,
redkina@spsl.nsc.ru, <https://orcid.org/0000-0002-3486-9711>*

Abstract. The robust, standardized metadata generated strictly under the scheme make research findings discoverable and understandable for users and information retrieval systems. Acquiring high-quality metadata for open access scientific resources and their harmonization with schemes and standards enhance visibility and citation of research findings, retrieval performance, proliferation of research, cooperation, and thematic structures studies. The goal of this study is to identify the spectrum of metadata and metadata schemes as reflected in the open access resources (ROA). For this purpose, the following problems were solved: 1) the list of fields of metadata, schemes and standards was examined, 2) the key elements of metadata for publications indexed in the largest ROA (Google Scholar, OpenAlex, Dimensions and Lens, and in the Russian Resources RISC and Cobalt) were identi-

fied. The author concludes on the prospects for the discovery resources and scientific information processes and on the need to consider evolving functionality, increasing discoverability of metadata accuracy and completeness. Elimination of record ambiguity with identifiers and record standards, linking metadata in different resources, introduction of extra metadata field to assess research results and to classify in the ROA classification, will enable relevant information retrieval and reduce document management in various information retrieval systems.

The study is accomplished within the framework of the RAS SB STLST Research Program "Development of scientific library functional model within open science information ecosystem" No. 122041100150-3.

Keywords: open access, resources, metadata, standards, schemes, bibliometric indicators, altmetrics, classification, indexing

Cite: Redkina N. S. Metadata and metadata schemes in open access research resources // Scientific and technical libraries. 2026. No. 5, pp. 146–163. <https://doi.org/10.33186/1027-3689-2026-5-146-163>

Введение

Мировой рынок информационных ресурсов трансформируется благодаря развитию экосистемы открытой науки и формированию ее культуры: разрабатываются инициативные программы, растет число сторонников, выступающих за открытие разнообразных результатов научных исследований; появляются новые РОД, которые отражают разные виды документов (статьи, патенты, наборы данных и др.), повышают доступность, увеличивают видимость и цитируемость научных работ, обеспечивают более быстрое распространение и оперативное применение новых идей и методов, повторное использование результатов, совместную работу и ускоренный обмен информацией, а также соблюдение требований финансирующих организаций, издательств, научных и образова-

тельных учреждений, стремящихся к соответствию принципам FAIR (Findable, Accessible, Interoperable, Re-usable)¹.

Об увеличении объема открытой информации свидетельствуют статистические данные Лейденского рейтинга 2024 г. (CWTS Leiden Ranking, <https://www.leidenranking.com>), в котором анализируются публикации за 2019–2022 гг. Нами было выявлено более 20 европейских университетов с долей публикаций открытого доступа свыше 90%. По данным ежегодного отчета cOAlition S средний мировой показатель открытого доступа в 2022–2024 гг. составил около 56% [1], а для участников cOAlition S – около 80% и выше.

Результатом инициативных проектов открытой науки становится все большее количество альтернативных коммерческим базам данных ресурсов [2–6], например: OpenAlex для поиска научной информации и библиометрического анализа среди более 271 млн документов; Lens – более 200 млн библиографических записей, собранных и гармонизированных из Microsoft Academic, PubMed и CrossRef, дополненной информацией открытого доступа OpenAlex и UnPaywall, а также ссылками на ORCID; Dimensions – метаданные около 160 млн публикаций и более 60 млн наборов данных в открытом доступе со встроенными альтернативными показателями. В этой ситуации возникает проблема не только поиска информационных ресурсов, но и развития их функционала, повышения качества и расширения элементов метаданных, способствующих нахождению релевантных работ среди миллионов других по различным параметрам.

Методика исследования

Цель исследования – определение полей и схем метаданных, отраженных в РОД. Для достижения цели был проведен контент-анализ публикаций, подтверждающих значимость стандартов и схем метаданных в развитии ресурсов. Далее были проанализированы поля метаданных

¹ FAIR (Findable, Accessible, Interoperable and Re-usable): Findable (находимость) – легкий поиск определенных наборов данных, Accessible (доступность) – удобный доступ (по условиям доступа и возможностям хранения в течение длительного времени), Interoperable (интероперабельность) – совместимость с другими наборами данных или программным обеспечением, Re-usable (повторное использование) – многократное (повторное) использование в дальнейших исследованиях.

в крупнейших научных РОД (OpenAlex, Dimensions, Google Scholar, РИНЦ, Lens, Cobalt), а также сервисные возможности и функционал ресурсов по использованию и выгрузке метаданных, широте охвата исследовательских элементов. Исследование направлено на выявление и анализ новых элементов метаданных (библиометрические и альтметрические показатели, предметные области, выделенные на основе собственных систем систематизации/классификации документов) и возможностей их использования для поиска релевантной информации, разностороннего количественного анализа и последующей визуализации. Результаты анализа демонстрируют необходимость использования нескольких баз данных для поиска и анализа документопотоков, несмотря на связи между РОД и использование метаданных из одинаковых источников.

Значение метаданных, особенности схем метаданных

Метаданные документов – структурированная или полуструктурированная информация, которая позволяет создавать и использовать документы, а также управлять ими в разное время в различных областях деятельности [7]. Политика управления метаданными не может быть эффективной, если они имеют низкое качество. Качество метаданных напрямую влияет на качество данных и, таким образом, на надежность их анализа и эффективность ресурса, а также позволяет оптимизировать время и стоимость обработки в среде больших данных, что является одной из самых больших проблем менеджеров данных [8].

Метаданные – это то, что индексируют поисковые системы (Google Scholar, Яндекс), библиотечные каталоги и агрегаторы, такие как BASE, CORE, а также иные РОД. Без богатых метаданных (ключевые слова, аннотация, название) статья или набор данных становятся «невидимыми» для научного сообщества. Метаданные и постоянные идентификаторы – это соединительная ткань, которая связывает воедино разнообразные результаты исследований, обеспечивая открытие, доступность и повторное использование. Несмотря на столь важную роль, текущая модель создания и обогащения метаданных чревата неэффективностью [9]. Обычно задача улучшения метаданных ложится исключительно на создателей или происходит в локальных ресурсах. Этот фрагментированный подход приводит к значительным разрывам в качестве и охвате,

создавая барьеры для бесшовной интеграции ресурсов, связывания данных и использования научной информации на разных платформах.

Схема метаданных – логический план, показывающий отношения между элементами метаданных, как правило, посредством установления правил использования и управления метаданными, особенно касающимися семантики, синтаксиса и степени обязательности данных [10]. При этом схемы метаданных документов должны быть отражены в форматах, которые обеспечивают возможность взаимодействия между системами, обмен информацией и миграцию, а также процессы передачи документов на хранение и уничтожение.

Среди наиболее часто упоминаемых форматов и схем метаданных выделим такие: как: UNIMARC (Universal Machine Readable Cataloging, универсальный MARC-формат), созданный для решения проблемы несовместимости между различными национальными MARC-форматами (Machine-readable cataloguing, «машиночитаемая каталогизация»); коммуникативный формат RUSMARC, представляющий собой адаптированный под государственные стандарты и правила каталогизации России формат UNIMARC; BIBFRAME (Bibliographic Framework, «библиографическая структура») – модель данных для библиографического описания, основанная на принципах связанных данных, разрабатываемая для замены MARC-форматов; Дублинское ядро (DC, Dublin Core) – набор элементов данных для описания разнообразных видов ресурсов (видео, изображений, веб-страниц и т. д.) и унификации метаданных в соответствии с двумя уровнями (простой, состоящий из 15 элементов и компетентный, включающий 18 элементов, в том числе указание на аудиторию, происхождение и правообладателя); MODS (Metadata Object Description Schema, Схема описания метаданных объектов), которая содержит 20 базовых элементов, уточняемых атрибутами, и обеспечивает возможность конвертации данных в Дублинское ядро и MARC-форматы.

Разные схемы оптимизированы для разных типов ресурсов. У каждой схемы свои особенности и характеристика, в литературе даются рекомендации относительно выбора и использования различных схем, например, описания веб-ресурсов [11] и др. Так, Схема описания объектов метаданных (MODS) предназначена для переноса выбранных данных из существующих записей MARC21, а также может быть использована для дополнения других форматов метаданных. Для некоторых приложе-

ний, особенно тех, которые используют записи MARC, будут преимуществами по сравнению с другими схемами метаданных, например, набор элементов шире, чем Dublin Core, и он больше совместим с библиотечными данными.

Наряду с универсальными схемами метаданных получают развитие специализированные стандарты в зависимости от дисциплины, типа данных, информационного ресурса и др. Так, Международной детской электронной библиотекой (International Children's Digital Library = ICDL) разработана схема метаданных, которая стремится отразить уникальное поведение ребенка при поиске информации для выбора книги [12]. Есть несколько элементов, которые ICDL использует как элементы метаданных: цвет обложки (переплета), рейтинг, чувство и др. Элемент «Чувство» обеспечивает информацию о тех эмоциях, которые может испытать ребенок, прочитавший книгу (радость, испуг, грусть или восхищение).

Для описания данных, связанных с охраной окружающей среды, используют язык экологических метаданных (Ecological Metadata Language, EML) [13], включая, к примеру, полные записи о времени и месте, методах отбора проб в полевых и лабораторных условиях, используемом оборудовании и программном обеспечении (марку, модель и версию, если применимо). EML включает поля метаданных для полной детализации статей с данными, которые публикуются в журналах, специализирующихся на обмене научными данными и их сохранении.

Наличие стандартов и схем метаданных способствует более эффективному обмену и совместному использованию данных, а создание дисциплинарных стандартов и рекомендаций ведет к формализации спецификаций метаданных. Широко используемый международный стандарт для описания, обмена или архивирования данных из социальных, поведенческих и экономических наук предлагается Альянсом инициатив по документированию данных – Data Documentation Initiative, DDI (<https://ddialliance.org/>). Для описания наборов данных, чтобы обеспечить их цитирование и воспроизводимость (включая поля «источник финансирования», «геолокация»), применяется специализированная схема DataCite Schema. В научных ПОД схемы представляют метаданные в структурированном формате (часто на XML, JSON, Дублинское ядро). Это основа для автоматизированного обмена данными между системами (например, между журналом и индексом цитирования), для создания API

и для сложного семантического поиска. Единые стандарты позволяют разным платформам (репозиториям, журналам, индексам) обмениваться информацией. Например, статья из arXiv может быть автоматически импортирована в институциональный репозиторий благодаря стандартным метаданным. Однако, хотя стандарты и есть, их много, и они не всегда совместимы друг с другом, это может создавать проблемы при использовании многочисленных РОД.

Современные схемы метаданных являются технической основой для выполнения FAIR-принципов. Специально подобранный информационный и образовательный ресурс по стандартам данных и метаданных, взаимосвязанный с базами данных и политиками в отношении данных, с возможностью поиска с использованием расширенной фильтрации (страна, типы объектов и др.) предлагается ресурсом FAIRsharing (<https://fairsharing.org>).

Развитие метаданных в информационных ресурсах

В библиотечном сообществе существует понимание того, что качественная работа с метаданными должна учитывать ожидания пользователей, чтобы облегчить обнаружение и использование ресурсов [14]. Метаданные превращают разрозненные файлы в управляемые и обнаруживаемые ресурсы. Метаданным отводится большое значение, они оцениваются Премией за качество [15], которая учреждена Crossref. Исследования демонстрируют, что доступность этих элементов метаданных со временем улучшилась, по крайней мере для журнальных статей, наиболее распространенного типа публикаций в Crossref, но многим издателям необходимо приложить дополнительные усилия для реализации полной открытости библиографических метаданных [16].

Метаданные с течением времени меняются. При этом форматы MARC, такие как MARC21 и RUSMARC, определяют структуру и правила записи данных, но не ограничивают количество полей. Результаты изучения более 6 млн записей книг с 1980 по 2018 г. в каталоге Библиотеки Конгресса показали, что в формате MARC используется более 200 уникальных полей и 1300 уникальных подполей метаданных, хотя большинство из них используется менее чем в 1% всех записей [17]. В то же время библиографические записи становятся все более сложными, с большим количеством полей и подполей на запись. Нельзя не со-

гласиться, что наличие поля метаданных бесполезно, если оно не заполнено данными [18]. Наиболее общепризнанные характеристики качественных метаданных: полнота, точность (устранение типографских ошибок, согласованное выражение личных имен и названий мест, использование стандартных сокращений), происхождение (соответствие стандартам), соответствие ожиданиям, логическая согласованность, своевременность и доступность. Отмечается и отсутствие аффилированности автора, атрибут языка, присутствие орфографических ошибок и др. [14, 19].

Инициативы открытой науки (Инициатива по открытому цитированию, Инициатива по открытым рефератам и др.), поддерживаемые научными издательствами по всему миру (MDPI, Wiley, др.), способствуют появлению новых элементов метаданных, облегчая исследователям поиск релевантной информации и помогая проводить разносторонний библиометрический анализ. Устоявшейся практикой стало включение аффилированности авторов в виде идентификаторов в Реестр исследовательских организаций (Research Organizations Registry [ROR]) и ORCID (Open Researcher and Contributor ID – Открытый идентификатор исследователя и сотрудника), аннотации и списка ссылок, что позволяет избегать ошибок в метаданных. В крупнейших научных информационных ресурсах включаются специальные поля: для сведений об авторе (имя автора, идентификатор исследователя, идентификатор ORCID, адрес электронной почты и т. д.), научной организации (сведения о городе, стране аффилированной организации исследователя/автора и т. д.), типе публикации (статья в журнале, материалы конференции, исправленный вариант, статья в печати и т. д.), цитированиях (общее количество цитирований, цитируемые ссылки, различные коэффициенты цитирования и т. д.), финансирующей организации (номер/идентификатор гранта и др.), классификации субъекта (например, область исследования и предметная область публикации) и т. д.

Другие часто используемые поля метаданных включают DOI, год публикации, язык публикации и т. д. Дополнительная информация, такая как метаданные о финансировании, связи между объектами (например, «является препринтом», «является обзором», «финансируется») и номера клинических испытаний (где это уместно), дает еще больше информации о том, как появился этот научный результат [20]. Идентификаторы гран-

тов и идентификаторы спонсоров, которые являются частью метаданных финансирования, делают связь между исследованием и его спонсорами явной. Исследование метрик на уровне статей, авторов, журналов, построение сети цитирований, изучение сотрудничества, ко-цитирования и т. д. – те возможности, которые предоставляются в некоторых уже ставших популярными научных ресурсах. Получение этих показателей требует наличия метаданных для публикации, патента и/или набора данных, индексированных в научных РОД.

Особенности метаданных в РОД

Рассмотрим особенности представления метаданных в РОД на примере научной публикации открытого доступа российского автора, опубликованной в журнале издательства Springer (*Redkina, N. S. The Library in the Information Ecosystem of Open Science / N. S. Redkina // Scientific and Technical Information Processing. 2021. Vol. 48, No. 4. P. 239–247. DOI 10.3103/S0147688221040043*) через призму некоторых метаданных (показатели цитирования, альтметрические показатели, классификационные системы) в различных РОД (OpenAlex, Dimensions, Lens, Google Scholar), а также на российских ресурсах РИНЦ и Cobalt.

Данные цитирования. Базы данных демонстрируют значительные различия в охвате цитирующей анализируемую статью публикаций: наиболее полные данные цитирования в Google Scholar (28), где индексируются публикации на разных языках (английский, русский, португальский, испанский, китайский) и даются пять источников отражения статьи (на сайте издательства, в eLIBRARY, PubMed Central, Europe PMC и специализированной БД ACM Digital Library). Почти идентичный охват в трех крупнейших РОД: OpenAlex (13), Dimension (10) и Lens (10). Специфика OpenAlex заключается в том, что учитываются ссылки из разных видов документов (статьи (8), препринты (2), обзоры (2), главы книг (1)) и др. РИНЦ, как и ожидалось, наиболее полно индексирует публикации на русском языке, а также некоторые международные журналы.

Показатель взвешенного по области исследования цитирования (Field-Weighted Citation Impact, FWCI) в OpenAlex 2.715 и процентиль цитирования 92.69 являются достаточно высокими. Это означает, что статья цитируется почти в три раза чаще, чем средняя публикация в той же области и за тот же год, и превосходит по цитированию 92,69% по-

добных статей. В Dimensions также, помимо общего показателя цитируемости работы, указываются недавние ссылки, коэффициент цитирования по полям и относительный коэффициент цитирования.

Представляют интерес библиометрические показатели изданий в Cobalt, в котором проиндексировано более 130 млн публикаций: показатель влияния или престижа научного журнала, SJR (SCImago Journal Rank, основан на данных из библиографической базы Scopus, издательство Elsevier)² – 0.178, CiteScore (среднее количество цитирований, полученных статьями научного журнала за определенный период) 0.9, Impact Factor 0.4. Кроме того, даются квартили SCImago и Web of Science и показатели цитирования из Dimensions. Эти значения характерны для специализированных журналов в области социальных наук и библиотечного дела, где показатели импакт-фактора традиционно ниже, чем в естественных науках.

Результаты анализа демонстрируют вариативность данных цитирования в разных библиографических базах. Наибольшую полноту охвата обеспечивает комбинация Google Scholar и OpenAlex, однако для специализированных исследований подключение тематических и региональных баз данных может значительно расширить результаты. В частности, для анализируемой статьи использование российской БД РИНЦ и отраслевой ACM Digital Library дало возможность получить дополнительные данные цитирования.

Альтметрические показатели (вовлеченность и общественный резонанс). Альтметрические показатели не учитываются в оценке результативности ученого, однако позволяют сделать анализ исследовательской деятельности более комплексно, измерить влияние и вовлеченность вокруг научных публикаций, учитывая упоминания, комментарии и обсуждения в социальных сетях или новостных сайтах, сохранения в закладках или количество просмотров и скачиваний на сайтах издательств или ресурсов.

Рассматриваемая публикация активно используется и сохраняется научным сообществом, но имеет ограниченный резонанс в социальных

² SJR > 1.0: журнал имеет влияние выше среднего в своей предметной категории. Значение 2.0 означает, что журнал вдвое влиятельнее среднего журнала в своей области. SJR = 1.0: влияние журнала является средним для его категории. SJR < 1.0: журнал имеет влияние ниже среднего в своей категории.

медиа. В Dimensions показано 49 читателей, которые загрузили статью в Mendeley – это высокий показатель, свидетельствующий о том, что статья сохранена для чтения и потенциального цитирования многими исследователями. Очень низкий общественный резонанс – всего 2 упоминания в X (Twitter). Это говорит о том, что обсуждение статьи практически не вышло за пределы академической среды. Альтметрические данные отражаются и на издательской платформе журнала Springer Nature, где зафиксировано 5021 доступ к статье, что также может свидетельствовать об интересе к данной проблематике.

Используемый в Dimensions сервис Altmetric рассчитывает оценку на основе внимания, которое статья получила в интернете. Анализируемая статья находится в 25-м процентиле (занимает 303 956-е место) из 432 732 отслеживаемых статей аналогичного возраста во всех журналах и в 1-м процентиле (занимает 1-е место) из одной отслеживаемой статьи аналогичного возраста в журнале Scientific and Technical Information Processing.

Согласно данным Lens, статья добавлена в три коллекции пользователей, связанные с исследовательскими практиками открытой науки и преподаванием. В РИНЦ публикация включена в восемь подборок пользователей, а также представлены данные о количестве просмотров, загрузок, оценках и отзывах.

Можно сделать вывод, что в РОД нет единых правил по представлению и учету альтметрических показателей.

Индексирование по предметным областям. Системы демонстрируют разные подходы к классификации статьи. Однотипная рубрика, связанная с библиотечным делом и информатикой, выделена в РИНЦ (ГРНТИ: Библиотечное дело. Библиотековедение; Рубрика OECD – Other social sciences); Dimensions (FoR: Research categories Fields of Research (ANZSRC 2020), 46 Information and Computing Sciences (Информационные и вычислительные науки)), 4610 Library and Information Studies (Библиотековедение и информационные исследования), а также на платформе издателя Springer Nature (Library Science, Information Studies).

Алгоритмы индексирования Lens определяют центральную тему – открытую науку. Вместе с тем включают дополнительные области исследования: Информатика, Экосистемы, Науки о данных, Всемирная паутина

на, Управление знаниями, Экология, Биология, математика и др. В Cobalt используется раздел: General Computer Science (Общая информатика).

В OpenAlex применяются собственные способы классификации результатов исследований, осуществленные с помощью искусственного интеллекта: темы, области знания и подполя (связанные с темами).

Тема для анализируемой статьи. Методы управления исследовательскими данными (этот блок статей посвящен практикам, проблемам и возможностям обмена данными и управления ими в научных исследованиях. Он охватывает такие темы, как открытая наука, управление исследовательскими данными, повторное использование данных, метаданные, экология, цифровые репозитории, принципы FAIR и цитирование данных).

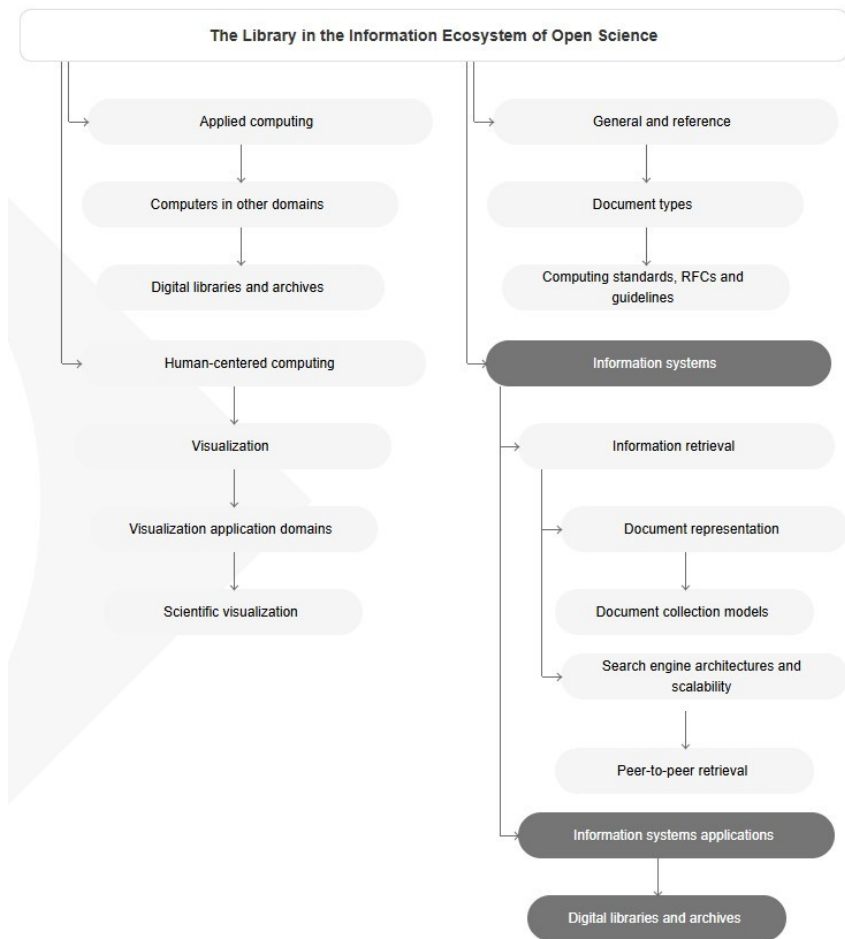
Подраздел. Информационные системы (академическое изучение систем с особым акцентом на информацию и дополнительные сети аппаратного и программного обеспечения, которые люди и организации используют для сбора, фильтрации, обработки, создания и распространения данных).

Область: Компьютерные науки.

Область: Физические науки (раздел естественных наук, изучающий неживые системы).

На издательской платформе приводятся только ключевые слова автора. Однако система предлагает осуществить поиск последних статей, книг и новостей по смежным темам, предложенным с использованием машинного обучения: Экосистемные услуги, Экосистемы, Издание Наука, Информационные исследования, Библиотекведение, Открытый исходный код.

Интересный подход к индексированию обнаружен в ACM Digital Library (см. рис.). Разделы областей знания были присвоены контенту посредством автоматической классификации и представлены в виде иерархии терминов: Information systems (Информационные системы), Information systems applications (Приложения информационных систем), Digital libraries and archives (Электронные библиотеки и архивы).



**Визуализация автоматической классификации
в ACM Digital Library статьи
The Library in the Information Ecosystem of Open Science**

Выводы

Метаданные помогают управлять цифровыми коллекциями: отслеживать версии, права доступа, форматы файлов и обеспечивать долгосрочную сохранность. Чем богаче и структурированнее метаданные, тем легче найти нужную работу, тем больше вероятность ее распространения в разнообразных РОД. Метаданные и схемы метаданных превращают разрозненные файлы и статьи в структурированную, взаимосвязанную и интеллектуальную экосистему знаний. Без них эффективный современный поиск, анализ и повторное использование научных результатов были бы невозможны, а система открытого доступа утонула бы в потоке неструктурированной информации. Создание качественных метаданных авторами, издателями и генераторами информационных ресурсов способствует эффективному распространению результатов исследований, их проверке и накоплению знаний в РОД.

Список источников

1. **Plan S:** Annual Review. 2024. URL: <https://www.coalition-s.org/wp-content/uploads/2025/04/Plan-S-2024-annual-review.pdf> (accessed: 05.11.2025).
2. **Шрайберг Я. Л.** Современные тенденции развития цифровизации общества: научно-образовательная и библиотечно-информационная среда : монография. Москва : ИНФРА-М, 2024. 664 с. DOI 10.12737/2155873.
3. **Шрайберг Я. Л., Вахрушев М. В., Гончаров М. В.** Открытый доступ: история, современное состояние и путь к открытой науке: учебник. Москва : Лань, 2020. 168 с.
4. **Редькина Н. С.** Ресурсы открытого доступа и инструменты для определения трендов развития науки // Научные и технические библиотеки. 2025; (4):90–113. DOI 10.33186/1027-3689-2025-4-90-113.
5. **Quirós D., Joaquina L., Ortega J. L.** Comparing Bibliographic Descriptions in Seven Free-Access Databases. 27th International Conference on Science, Technology and Innovation Indicators (STI 2023). 2023. DOI 10.55835/6436c590b3340c364be5b2c7.
6. **Alperin J. P., Portenoy J., Demes K. [et al.]** An analysis of the suitability of OpenAlex for bibliometric analyses. arXiv. 2024. April 26. URL: <https://arxiv.org/abs/2404.17663> (accessed: 03.10.2025). DOI 10.48550/arXiv.2404.17663.
7. **ГОСТ Р ИСО 15489-1-2019.** Национальный стандарт Российской Федерации. Система стандартов по информации, библиотечному и издательскому делу. Информация и документация. Управление документами. Часть 1. Понятия и принципы (утв. Приказом Росстандарта от 26.03.2019 N 101-ст). Москва, Стандартинформ, 2019. 24 с.

8. **Elouataoui W., El Alaoui I., Gahi Y.** Metadata Quality Dimensions for Big Data Use Cases. Proceedings of the 2nd International Conference on Big Data, Modelling and Machine Learning. 2021. 1: 488-495. DOI 10.5220/0010737400003101.
9. **Buttrick A., Praetzellis M.** Advancing Metadata Quality: An Open Call to Collaborate / FORCE 11. 2024. October 8. DOI 10.54900/y7yyq-2rw17.
10. **ГОСТ Р ИСО 15489-1-2019.** Национальный стандарт Российской Федерации. Система стандартов по информации, библиотечному и издательскому делу. Информация и документация. Управление документами. Часть 1. Москва, Стандартинформ, 2019. 28 с.
11. **Негуляев Е. А.** Универсальные схемы метаданных и задача описания веб-ресурсов // Библиосфера. 2006. 1:43-54.
12. **Бик Дж., Олсон Х. Э.** Сравнение схем метаданных: AACR2+ или схема метаданных ICDL. URL: <https://www.ifla.org/past-wlic/2011/80-beak-ru.pdf> (дата обращения: 05.11.2025).
13. **Jones M. B., O'Brien M., Mecum B. [et al.]** Ecological Metadata Language version 2.2.0. KNB Data Repository. 2019. DOI 10.5063/F11834T2.
14. **Shi J., Nason M., Tullney M., Alperin J.** Identifying Metadata Quality Issues Across Cultures. College & Research Libraries. 2025. 86 (1): 101. DOI 10.5860/crl.86.1.101.
15. **Korzec K., Tkaczyk D.** Meet Six Winners of the First Ever Crossref Metadata Awards / Metadata Blog. May 7, 2025. DOI 10.64000/xh94q-w7335.
16. **Van Eck N. J., Waltman L.** Crossref as a Source of Open Bibliographic Metadata. MetaArXiv. 2025. May 12. DOI 10.31222/osf.io/smxe5_v2.
17. **Li K., Dobreski B., Busch M.** How are library cataloging metadata used differently over time: a large-scale quantitative analysis of the Library of Congress catalog. Journal of Documentation. 2025. 81(2): 385-402. DOI 10.1108/JD-08-2024-0199.
18. **Lide D. R., Kehiaian H. V.** CRC handbook of thermophysical and thermochemical data. CRC Press. 2020. 518 p.
19. **Barnett A.** Spelling Errors in Health and Medical Abstracts: An Observational Analysis from 2008 to 2024. OSF Preprints. 2025. July 29. DOI 10.31219/osf.io/v8n9q_v2.
20. **Amdekar M. S.** Scholarly Metadata as Trust Signals: Opportunities for Journal Editors. Science Editor. 2024. November 6. DOI 10.36591/se-4704-10.

References

1. **Plan S:** Annual Review. 2024. URL: <https://www.coalition-s.org/wp-content/uploads/2025/04/Plan-S-2024-annual-review.pdf> (accessed: 05.11.2025).
2. **Shrai'berg Ia. L.** Sovremenny'e tendentsii razvitiia tcifrovizatscii obshchestva: nauchno-obrazovatel'naia i bibliotechno-informatcionnaia sreda : monografiia. Moskva : INFRA-M, 2024. 664 s. DOI 10.12737/2155873.
3. **Shrai'berg Ia. L., Vakhrushev M. V., Goncharov M. V.** Otkry'ty'i dostup: istoriia, sovremennoe sostoiianie i put`k otkry'toi` nauke: uchebник. Moskva : Lan`, 2020. 168 s.
4. **Red'kina N. S.** Resursy` otkry'togo dostupa i instrumenty` dlia opredeleniia trendov razvitiia nauki // Nauchny'e i tekhnicheskie biblioteki. 2025; (4):90–113. DOI 10.33186/1027-3689-2025-4-90-113.

5. **Quirós D., Joaquina L., Ortega J. L.** Comparing Bibliographic Descriptions in Seven Free-Access Databases. 27th International Conference on Science, Technology and Innovation Indicators (STI 2023). 2023. DOI 10.55835/6436c590b3340c364be5b2c7.
6. **Alperin J. P., Portenoy J., Demes K. [et al.]** An analysis of the suitability of OpenAlex for bibliometric analyses. arXiv. 2024. April 26. URL: <https://arxiv.org/abs/2404.17663> (accessed: 03.10.2025). DOI 10.48550/arXiv.2404.17663.
7. **GOST R ISO 15489-1-2019.** Nacional'ny'i standart Rossiiskoi` Federatsii. Sistema standartov po informatsii, biblioteknomu i izdatel'skomu delu. Informatsiia i dokumentatsiia. Upravlenie dokumentami. Chast' 1. Poniatii i printsiy` (utv. Prikazom Rosstandarta ot 26.03.2019 N 101-st). Moskva, Standartinform, 2019. 24 s.
8. **Elouataoui W., El Alaoui I., Gahi Y.** Metadata Quality Dimensions for Big Data Use Cases. Proceedings of the 2nd International Conference on Big Data, Modelling and Machine Learning. 2021. 1: 488-495. DOI 10.5220/0010737400003101.
9. **Buttrick A., Praetzelis M.** Advancing Metadata Quality: An Open Call to Collaborate / FORCE 11. 2024. October 8. DOI 10.54900/y7yyq-2rw17.
10. **GOST R ISO 15489-1-2019.** Nacional'ny'i standart Rossiiskoi` Federatsii. Sistema standartov po informatsii, biblioteknomu i izdatel'skomu delu. Informatsiia i dokumentatsiia. Upravlenie dokumentami. Chast' 1. Moskva, Standartinform, 2019. 28 s.
11. **Neguliaev E. A.** Universal'ny'e skhemy` metadanny`kh i zadacha opisaniia veb-resursov // Bibliosfera. 2006. 1:43-54.
12. **Bik Dzh., Olson KH. E`.** Sravnenie skhem metadanny`kh: AACR2+ ili skhema metadanny`kh ICOL. URL: <https://www.ifla.org/past-wlic/2011/80-beak-ru.pdf> (data obrashcheniia: 05.11.2025).
13. **Jones M. B., O'Brien M., Mecum B. [et al.]** Ecological Metadata Language version 2.2.0. KNB Data Repository. 2019. DOI 10.5063/F11834T2.
14. **Shi J., Nason M., Tullney M., Alperin J.** Identifying Metadata Quality Issues Across Cultures. College & Research Libraries. 2025. 86 (1): 101. DOI 10.5860/crl.86.1.101.
15. **Korzec K., Tkaczyk D.** Meet Six Winners of the First Ever Crossref Metadata Awards / Metadata Blog. May 7, 2025. DOI 10.64000/xh94q-w7335.
16. **Van Eck N. J., Waltman L.** Crossref as a Source of Open Bibliographic Metadata. MetaArXiv. 2025. May 12. DOI 10.31222/osf.io/smxe5_v2.
17. **Li K., Dobreski B., Busch M.** How are library cataloging metadata used differently over time: a large-scale quantitative analysis of the Library of Congress catalog. Journal of Documentation. 2025. 81(2): 385-402. DOI 10.1108/JD-08-2024-0199.
18. **Lide D. R., Kehiaian H. V.** CRC handbook of thermophysical and thermochemical data. CRC Press. 2020. 518 p.
19. **Barnett A.** Spelling Errors in Health and Medical Abstracts: An Observational Analysis from 2008 to 2024. OSF Preprints. 2025. July 29. DOI 10.31219/osf.io/v8n9q_v2.
20. **Amdekar M. S.** Scholarly Metadata as Trust Signals: Opportunities for Journal Editors. Science Editor. 2024. November 6. DOI 10.36591/se-4704-10.

Информация об авторе / Author

Редкина Наталья Степановна –
доктор пед. наук, главный научный
сотрудник, заведующий Отделом
научных исследований открытой
науки ГПНТБ СО РАН, Новосибирск,
Российская Федерация
redkina@spsl.nsc.ru

Natalya S. Redkina – Dr. Sc.
(Pedagogy), Chief Researcher, Head,
Open Science Studies, SB RAS State
Public Scientific and Technological
Library, Novosibirsk,
Russian Federation
redkina@spsl.nsc.ru