СОВРЕМЕННЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

УДК 001.103.2:025.32 + 002.1 – 021.341
https://doi.org/10.33186/1027-3689-2025-1-98-119

Проблемы полноты и корректности ввода метаданных в репозиториях российских образовательных организаций

А. А. Стукалова

ГПНТБ СО РАН, Новосибирск, Российская Федерация,
Stukalova@gpntbsib.ru, https://orcid.org/0000-0003-2202-943X

Аннотация. Большую роль в структуре открытой науки играют институциональные репозитории (ИР), предоставляющие свободный доступ к научным и образовательным коллекциям. Одним из важнейших компонентов, отвечающих за эффективность их функционирования, являются метаданные, обеспечивающие видимость контента ИР.

Чтобы определить детальность и корректность заполнения элементов метаданных объектов ИР, проанализированы 25 репозиториев российских образовательных организаций – участников программы «Приоритет-2030». В результате проведённого исследования выявлено, что метаданные содержат в основном базовые элементы, включающие описательную информацию: заглавие, выходные данные, язык и тип документа. Значительно реже в них содержатся сведения о соавторах, редакторах, коллективах, источниках публикаций. Структурные данные, включающие сведения о составе, форматах, объёме документа, также отражены не во всех метаданных объектов ИР. В редких случаях в них отражены связи, охват документа, сведения о правах доступа.

При этом от полноты предоставления метаданных зависит эффективность поиска в репозиториях. С их помощью реализуются поисковые возможности, обеспечиваются уточнение результатов поиска, его сортировка. Чем больше элементов включают метаданные, тем детальнее представлена информация об объектах репозитория и тем выше вероятность того, что эта информация будет найдена пользователем.

Ключевые слова: открытый доступ, институциональный репозиторий, метаданные, Дублинское ядро, классификационные индексы

Благодарности: статья подготовлена в рамках реализации научного проекта ГПНТБ СО РАН (2021–2026 гг.) «Разработка модели функционирования научной библиотеки в информационной экосистеме открытой науки» № 122041100150-3.

Для цитирования: Стукалова А. А. Проблемы полноты и корректности ввода метаданных в репозиториях российских образовательных организаций // Научные и технические библиотеки. 2025. № 1. С. 98–119. https://doi.org/10.33186/1027-3689-2025-1-98-119

MODERN INFORMATION TECHNOLOGIES

UDC 001.103.2:025.32 + 002.1 – 021.341
https://doi.org/10.33186/1027-3689-2025-1-98-119

On the completeness and correctness of input metadata in the repositories of Russian educational organizations

Anna A. Stukalova

State Public Scientific and Technological Library of the Russian Academy of Sciences Siberian Branch, Novosibirsk, Russian Federation, Stukalova@gpntbsib.ru, https://orcid.org/0000-0003-2202-943X

Abstract. The institutional repositories (IR) with the open access to research and educational collections play significant role in the open science infrastructure. The metadata provide IR content discoverability and repository effectiveness. The author analyzed 25 repositories of Russian educational organization to specify the metadata detail level and correctness. The educational organizations under the study are the members of “The Priority-2030” Program. It was found that the metadata comprised mainly the basic elements: title, publication data, document language and type while the data on co-authors, editors, research teams and sources are much more rare. The structural data, i.e. data on contents, formats, volume are often not included, too. Very rarely, the metadata comprise links, subject scope, and accessibility information.

However, the metadata are important for information retrieval efficiency, search capabilities, search results refinement, sorting. The more elements the metadata comprise, the more details of repository objects are provided, the higher information discoverability is.

Keywords: open access, institutional repository, metadata, Dublin Core, classification number

Acknowledgements: The article is prepared within the framework of the SB RAS State Public Scientific and Technological Library's research project "Development of the scientific library model for the open science information ecosystem" No. 122041100150-3.

Cite: Stukalova A. A. On the completeness and correctness of input metadata in the repositories of Russian educational organizations // Scientific and technical libraries. 2025. No. 1, pp. 98–119. https://doi.org/10.33186/1027-3689-2025-1-98-119

В России продолжают активно развиваться институциональные репозитории (ИР), предоставляющие открытый доступ (ОД) к результатам научных исследований. Проблемам и преимуществам их использования посвящено большое количество публикаций зарубежных [**1–5**] и отечественных авторов [**6–8**]. Большое внимание уделяется вопросам программного обеспечения для репозиториев [**9**, **10**], проблемам обмена данными [**11**], доступности репозиториев, развитию их функциональных возможностей [**12**]. Однако использованию метаданных в ИР посвящено небольшое количество публикаций [**13**, **14**].

В 2020 г. сотрудники НЭИКОН опубликовали методические рекомендации, регламентирующие единые правила формирования метаданных с целью их передачи в Национальный агрегатор открытых репозиториев (НОРА), а также в зарубежные агрегаторы. В рекомендациях предлагается сделать обязательными шесть полей. Но НОРА не контролирует обязательность их заполнения на входе. Если они отсутствуют в записях, материалы всё равно будут загружены в агрегатор [**15**].

В то же время в каждом ИР продолжают применяться свои правила описания ресурсов. Отсутствие единообразия при вводе метаданных препятствует эффективности поиска необходимых пользователю научных исследований, интеграции репозиториев с внешними системами.

Определение термина «метаданные»

Метаданные являются упорядоченной информацией, которая всесторонне описывает контент и его основные элементы. От того, насколько аккуратно и тщательно будет описан каждый объект, зависит эффективность использования цифровой коллекции [**16**. С. 71].

В настоящее время существует большое количество определений метаданных. Термин «metadata» состоит из двух терминов – «meta» и «data», первый из которых может быть переведён как «абстрактный», а второй – «данные» [**17**. С. 52].

В аспекте истории информатики понятие «метаданные» является интегральным по отношению к таким традиционным понятиям, как «форматы представления данных», «языки описания данных», «лингвистическое обеспечение автоматизированных информационных систем». Главное отличие «метаданных» от перечисленных понятий заключается в более общем характере, подобно тому, как цифровой объект (информационный ресурс) является более широким понятием по сравнению с понятиями «документ» или «единица хранения предыдущих поколений информационных систем» [**18**. С. 26].

Согласно ГОСТ Р 55750-2013, метаданные электронного образовательного ресурса – это структурированные данные, определяющие основные характеристики электронных образовательных ресурсов и предназначенные для их систематизации и эффективного применения в электронных информационно-образовательных средах [**19**. С. 2].

Подробно понятие «метаданные» проанализировано в статьях Т. Ф. Берестовой [**20**], Н. А. Мазова, О. Л. Жижимова [**17**].

В широком понимании метаданные – это данные, описывающие содержание и структуру ресурса [**15**. С. 4].

Большинство исследователей выделяет следующие типы метаданных:

описательные (библиографическая информация или другие сведения о содержании цифровых объектов);

административные (включают сведения о правах, разрешениях на доступ, коррекцию данных, данные о пользователе, технологические данные и др.);

структурные (включают сведения о структуре, форматах, объёме и других элементах цифрового произведения).

Отдельно выделяют идентификатор – как тип метаданных, который представляет цифровой объект для внешних пользователей и различных приложений [**18**. С. 27].

Для описания ресурсов применяются различные системы метаданных, которые подробно рассмотрены в статьях Е. А. Негуляева [**21**] и А. Б. Антопольского [**18**]. Наиболее распространённой является система метаданных Дублинского ядра – Dublin Core (DC), разработанного в рамках движения ОД. Преимуществом использования DC является возможность создания такого механизма, который при минимальных затратах на описание позволяет искать и находить данные вне зависимости от языковой принадлежности, тематики и места их нахождения [**22**].

В 2019 г. вышел ГОСТ Р 7.0.10-2019, характеризующий 15 элементов DC, которые составляют так называемый набор метаданных [**23**]. Именно он является основой для обмена информацией между репозиториями и интеграции их контента в агрегаторах. Этот набор позволяет описать объекты на довольно общем уровне, не позволяющем детализировать важные дополнительные аспекты [**15**. С. 6]. Ещё одной проблемой являются широкие формулировки этих элементов, позволяющие интерпретировать их заполнение по своему усмотрению.

Расширенный (квалифицированный) набор метаданных содержит 50+ термов, включающих 15 основных элементов, и термы, детализирующие семантику базового набора [**24**]. Восемь элементов могут быть уточнены с помощью одного или более квалификаторов (например, title, alternative), а набор разрешённых схем кодирования (определяемых необязательным атрибутом scheme) позволяет создавать квалификаторы для 10 элементов [**25**. С. 180].

Метаданные, как правило, содержат элементы, по которым обычно ведут поиск пользователи. С их помощью обеспечиваются сортировка, уточнение результатов поиска. Поэтому, чем больше элементов включают метаданные, тем детальнее представлена информация об объектах ИР и тем выше вероятность того, что эта информация будет найдена пользователем. В то же время большое количество метаданных не

гарантирует видимость описываемой информации, так как не всегда эти сведения введены корректно.

С целью определения полноты и корректности ввода метаданных ИР были изучены объекты 25 репозиториев отечественных образовательных организаций – участников программы «Приоритет-2030» на предмет заполнения основных элементов DC. В каждом ИР проанализированы метаданные 50 документов различных типов. В основном: монографии, выпускные квалификационные работы (ВКР), авторефераты и диссертации, статьи из журналов и сборников, материалы конференций, препринты. В некоторых ИР рассмотрены и другие типы. Например, электронные архивы Севастопольского государственного университета и Белгородского государственного национального исследовательского университета содержат презентации; репозиторий Российского университета дружбы народов и электронная библиотека (ЭБ) Научно-технической библиотеки Томского политехнического университета – патенты; архив электронных ресурсов Сибирского федерального университета – презентации, объекты авторского права, газеты; репозиторий Томского государственного университета – газеты, изоиздания, карты, онлайн-курсы; каталог научно-образовательных ресурсов Казанского федерального университета – экспонаты выставок; цифровой репозиторий Южного федерального университета – презентации, студенческие работы, наборы данных, патенты.

К сожалению, только в 14 ИР дана возможность ознакомиться с полными описаниями метаданных. В остальных для удалённого пользователя представлена только краткая запись, в которой не указаны метки элементов и кодированная информация. Например, в ИР Башкирского государственного медицинского университета дана возможность ознакомиться с краткой записью, содержащей основные сведения о документе:

СТАНОВЛЕНИЕ РОССИЙСКОЙ ОНКОЛОГИЧЕСКОЙ СЛУЖБЫ ПОСЛЕ ВЕЛИКОЙ ОТЕЧЕСТВЕННОЙ ВОЙНЫ

ESTABLISHMENT OF THE RUSSIAN ONCOLOGY SERVICE AFTER THE GREAT PATRIOTIC WAR

Ш. Х. Ганцев, Sh. Kh. Gantsev, А. М. Беляев, A. M. Belyaev, В. В. Старинский, V. V. Starinsky, Э. Р. Иванова, E. R. Ivanova, Ш. Р. Кзыргалин, Sh. R. Kzyrgalin (Креативная хирургия и онкология, № 2, 2020)

В то время как в других ИР (как правило, использующих программное обеспечение DSpace) возможно просмотреть как краткие описания, так и полные записи метаданных. Например, запись электронного архива ОД Белгородского государственного технического университета им. В. Г. Шухова выглядит следующим образом:

Поле DC Значение Язык

dc.contributor.author Семикопенко, И. А. –

dc.contributor.author Воронов, В. П. –

dc.contributor.author Вялых, С. В –

dc.contributor.author Маняхин, А. С. –

dc.date.accessioned 2018-04-13T06:40:01Z –

dc.date.available 2018-04-13T06:40:01Z –

dc.date.issued 2018 –

dc.identifier.uri http://dspace.bstu.ru/jspui/handle/123456789/101 –

dc.description К вопросу о пропускной способности узла предварительного измельчения и классификации материала в дезинтеграторе / И. А. Семикопенко, В. П. Воронов, С. В. Вялых, А. С. Маняхин // Вестник БГТУ им. В. Г. Шухова. 2018. No 4. С. 119–123.ru_RU

dc.description.abstract Одним из недостатков работы дезинтеграторов является незначительное количество соударений частиц материала в камере помола. В связи с этим была создана дезинтеграторная установка, в которой осуществляется предварительное измельчение материала при его загрузке в камеру помола. Наряду с измельчением в зоне загрузки осуществляется непрерывный отвод части измельчённого материала. В данной статье получено аналитическое выражение, позволяющее определить пропускную способность узла предварительного измельчения и классификации материала в дезинтеграторе. Представлена схема для определения пропускной способности данного узла с учётом непрерывного отвода мелких частиц через призматические каналы верхней ступени многоступенчатого корпуса. Анализ полученных соотношений позволяет сделать вывод о том, что пропускная способность узла предварительного измельчения и классификации материала и количество отделяемого продукта зависят от геометрических и технологических параметров.

dc.language.iso other ru_RU

dc.publisherИздательство БГТУ им. В. Г. Шухова ru_RU

dc.subject Авторы БГТУ ru_RU

dc.subject многоступенчатый корпус ru_RU
dc.subject узел ru_RU
dc.subject частица ru_RU
dc.subject пропускная способность ru_RU
dc.title К вопросу о пропускной способности узла предварительного измельчения и классификации материала в дезинтеграторе ru_RU
dc.type Article ru_RU
Располагается в коллекциях: 2018 год

Результаты анализа метаданных репозиториев образовательных организаций

Анализ метаданных объектов ИР дал следующие результаты:

Title/Заглавие. Согласно ГОСТ Р 7.0.10-2019 этот элемент содержит имя, данное ресурсу. Это может быть человек, организация или служба (сервис). Заглавие – один из основных элементов метаданных, необходимость заполнения которого не вызывает сомнений. Анализ метаданных показал, что во всех объектах изученных ИР указано название документа. В электронных архивах Сибирского государственного медицинского университета, Новосибирского государственного университета, репозиториях Томского государственного университета, Российского университета дружбы народов, помимо основного названия, в элементе dc.title.alternative дано альтернативное название на английском языке.

Creator/Создатель. В данном элементе следует указывать информацию о лице, несущем первичную ответственность за создание ресурса. В метаданных на текстовые материал указывается автор, на электронные ресурсы – исполнители, фотографы или иллюстраторы, композиторы и т. п. Анализ метаданных показал, что сведения об авторах, независимо от их количества и местоположения в перечне всех лиц, участвовавших в создании того или иного объекта, указаны в элементе метаданных contributor/соисполнитель, а не в creator/создатель.

Contributor/Соисполнитель. Согласно ГОСТ Р 7.0.10-2019 элемент включает сведения о лице, ответственном за вклад в создание ресурса. Соисполнителем может быть человек, организация или служба (сервис), которые не являются авторами (не обозначены в элементе автор), но внесли интеллектуальный вклад в ресурс, причём этот вклад вторичен по отношению к любому человеку или организации, указанной в числе

авторов (например, редактор, переводчик, иллюстратор) [**18**]. Проведённое исследование показало, что данный элемент присутствует в метаданных объектов ИР. В этом элементе указаны не только соисполнители, но и основной исполнитель – первый автор публикаций. В большинстве изученных ИР данный элемент продублирован по количеству соисполнителей, то есть имеет отдельное повторение для каждого автора, редактора и т. п. Однако в некоторых ИР в элементе dc.contributor.author указан только первый исполнитель, хотя документ создан несколькими авторами. Это приводит к тому, что при поиске по соавторам данный документ не будет найден пользователем. В электронном архиве Севастопольского государственного университета в элементе dc.contributor.author указаны не только ФИО автора, но и пояснения о его месте работы и должности. Помимо сведений об авторах, в метаданных объектов репозиториев Томского государственного университета, Тольяттинского государственного университета, архива электронных ресурсов Сибирского федерального университета даны сведения об организациях.

Date/Даты. Данный элемент содержит информацию о моменте или отрезке времени, связанном с событием жизненного цикла ресурса. Может обозначать время на любом уровне дробности. Это ещё один элемент, не вызывающий вопросов по заполнению. Во всех изученных ИР в метаданных указана дата публикации (год). В большинстве ИР этот элемент дублируется на квалификаторы dc.date.accessioned (дата создания записи в системе), dc.date.available (дата доступности), dc.date.issued (дата издания). В электронном научном архиве Уральского федерального университета помимо перечисленных дат указаны даты проведения конференций, семинаров.

Publisher/Издатель. Согласно ГОСТ Р 7.0.10-2019 элемент должен включать информацию о лице, ответственном за предоставление ресурса пользователям. Может быть человек, организация или служба (сервис) (например, издательский дом, университет или общество). Проведённый анализ показал, что большинство метаданных объектов ИР включает этот элемент. В некоторых из них помимо информации об организациях, отвечающих за предоставление ресурса, даны сведения о городе, в котором данная организация размещена. Следует обратить внимание, что в некоторых метаданных ИР эти сведения введены с сокращением слов, например: Изд-во Самар. ун-та. Это негативно вли-

яет на поисковые возможности ИР, так как ни один пользователь не будет искать информацию в сокращённой форме.

Identifier/Идентификатор. Неразрывная строка символов, идентифицирующая ресурс, организацию или лицо самостоятельно или в сочетании с другими элементами. В большинстве случаев в данных элементах указаны URI[1], ISBN/ISSN, DOI в соответствующих элементах dc.identifier.isbn, dc.identifier.uri, dc.identifier.doi. URI используется как ссылка, ведущая на ресурс в репозитории. Во многих ИР указано библиографическое описание документа в элементе dc.identifier.citation. В репозитории Южного федерального университета помимо DOI и ISBN указан трек-номер. В изученных ИР этот элемент, как правило, заполнен корректно, но обнаружены примеры, в которых были введены несоответствующие данные. Например, в элементе dc.identifier.citation введена фраза «выходные данные».

Type/Тип ресурса. Согласно ГОСТ Р 7.0.10-2019 элемент должен отражать информацию о природе или жанре содержания ресурса. Анализ объектов показал, что данный элемент присутствует в метаданных объектов ИР. В основном в нём введены сведения о типе документа: монография, научная статья, материалы конференции и т. п. В некоторых ИР это поле заполнено на английском языке.

Format/Формат. Указывается формат файла, физический носитель или размеры ресурса. Например, в электронном научном архиве Уральского федерального университета указаны размер 5,02 MB, формат Adobe PDF. Данный элемент был найден в метаданных 50% изученных ИР, так как в кратких записях эта информация, как правило, не отражена.

Source/Источник. Согласно ГОСТ Р 7.0.10-2019 элемент отражает информацию о связанном (родственном) ресурсе, послужившим источником данных для описываемого ресурса. Анализ метаданных показал, что в этом элементе указаны названия периодических изданий, в которых опубликована статья. В метаданных на главы из книг ИР «Публикации» Национального исследовательского университета «Высшая школа экономики»» указано название книги, откуда взяты эти главы. Однако в 38% ИР элемент dc.source не заполнен, соответственно в ре-

---

[1] URI (Uniform Resource Identifier – единый идентификатор ресурса) – неразрывная строка символов, идентифицирующая абстрактный или физический ресурс.

зультате поиска по названию журнала сведений об этих объектах не будет.

Language/Язык. Данный элемент содержит язык содержательной части ресурса. Анализ метаданных показал, что эта информация представлена во всех записях ИР с полным описанием метаданных. К сожалению, убедиться в наличии данного элемента в кратких записях не представляется возможным, так как в них эта информация, как правило, не видна.

Relation/Связь. Связанный (родственный) ресурс. Рекомендуется указывать родственный ресурс посредством URI или другой строкой символов в соответствии с какой-либо формальной системой идентификации. Проведённый анализ показал, что данный элемент не заполняется в метаданных объектов ИР. Исключением является Электронный научный архив Уральского федерального университета. В элементе метаданных dc.relation.ispartof указаны издания, в которых опубликован материал. В электронном архиве Сибирского государственного медицинского университета в элементе dc.relation.ispartofseries введено название серии. В ИР Томского государственного университета при просмотре ссылки «связи документа» открывается окно с названиями организаций, кафедр, журналов, коллекций, к которым относится просматриваемой документ.

Coverage/Охват. ГОСТ Р 7.0.10-2019 даёт довольно широкое описание данного элемента. Это характеристика содержания ресурса в пространстве или времени, пространственная применимость ресурса или юрисдикция, под которой ресурс имеет силу. Пространственной характеристикой тематики или применимости может быть имя места или указание географических координат. Время тематики ресурса может быть задано наименованием периода, датой или диапазоном дат. Юрисдикцией может быть наименование административной единицы или географического места, к которому применяется ресурс.

Анализ метаданных объектов ИР показал, что этот элемент почти не применяется. Исключением является ИР «Публикации» Национального исследовательского университета «Высшая школа экономики», в метаданных которого указано научное направление. В репозитории Самарского университета в элементе метаданных dc.coverage.spatial указаны ключевые слова.

Rights/Права доступа. Данный элемент должен включать сведения о правах на использование и управление ресурсом. Эти сведения необходимы для удалённого пользователя, поскольку они информируют о том, как можно получить доступ к полному контенту. Сотрудниками НЭИКОН озвучены следующие допустимые значения для этого элемента:

open access (ОД) – свободный доступ любых пользователей к метаданным и полным текстам;

restricted access (доступ с эмбарго) – наличие ограничений в доступе, например доступ только для сотрудников организации;

embargoed access (ограниченный доступ) – отложенный доступ к полным текстам (после периода эмбарго);

metadata only access (доступ только к метаданным) – доступ к метаданным, но не к полному тексту [**24**].

Анализ метаданных объектов репозиториев показал, что этот элемент практически не заполняется. Исключением является репозиторий Томского государственного университета, в котором указан доступ ресурса на сайте издателя, либо в сети университета. В цифровом репозитории Южного федерального университета указаны следующие разрешения на доступ:

на просмотр данной страницы: разрешено всем пользователям, включая неавторизованных;

на сохранение файлов материала: не разрешено никому;

на чтение файлов материала в браузере: не разрешено никому.

Description/Описание. Согласно ГОСТ Р 7.0.10-2019 данный элемент включает сведения о ресурсе, которые могут содержать реферат, аннотацию, оглавление, графическое представление, текстовое изложение. Анализ метаданных ИР показал, что только в трёх из них не предоставляются аннотации или рефераты. В репозиториях Южного федерального университета и Самарского университета, помимо аннотации, метаданные включают отрывок полного текста. В репозитории Российского университета дружбы народов и в электронном архиве Новосибирского государственного университета в элементе dc.description.abstract аннотации продублированы на английском языке, а в элементе dc.description.sponsorship представлена информация о гранте, при поддержке которого издана книга. В ИР «Публикации» Национального исследовательского университета «Высшая школа эко-

номики» указано название проекта, по результатам которого подготовлена публикация.

Subject/Предмет. Одним из значимых элементов, отвечающих за эффективность предметного поиска в ИР, является элемент subject, который отражает тематику ресурса, его предметное содержание. Данный элемент выражается ключевыми словами, ключевыми фразами или классификационными кодами. Проведённый анализ показал, что в 75% метаданных этот элемент заполнен. В большинстве ИР он содержит ключевые слова, при этом элемент, как правило, имеет множество повторений. Например, в электронном архиве Сибирского государственного медицинского университета метаданные одного из объектов включали 14 повторений поля subject, subject_mesh. Довольно редко в репозиториях применяются классификационные коды. Например, в репозиториях Южного федерального университета, Института вулканологии и сейсмологии помимо ключевых слов указаны коды ГРНТИ. В метаданных репозитория Тверского госуниверситета и архива электронных ресурсов Сибирского федерального университета содержится УДК. Коды УДК, ГРНТИ присутствуют в метаданных репозитория Самарского университета. В элементах subject метаописаний ЭБ Научно-технической библиотеки Томского политехнического университета им. В. А. Обручева содержатся ключевые слова, предметные рубрики, ББК либо УДК. Кроме того, в записях на авторефераты диссертаций и диссертации указан общероссийский классификатор специальностей высшей научной квалификации (ОКСВНК). В ЭБ Алтайского государственного университета в метаданных на ВКР указаны направления подготовки. Например, 09.03.01 – Информатика и вычислительная техника.

Предоставление возможности поиска по перечисленным данным значительно повышает точность тематического поиска.

Специфические данные. Помимо рассмотренных элементов метаданных в записях репозиториев были найдены и другие элементы, характерные для заполнения различных типов изданий. Например, в метаданных на патенты цифрового репозитория Южного федерального университета указан номер патента; в репозитории Российского университета дружбы народов указаны номер патента, дата регистрации, дата начала срока, база данных, тип, направления исследования, статус, номер и дата заявки; в ЭБ Научно-технической библиотеки Томского

политехнического университета в области специфических требований указаны номер патента, дата регистрации, дата начала срока.

В метаданных на авторефераты диссертаций и диссертации репозитория Российского университета дружбы народов указаны учёная степень, место защиты, специальность, а в репозитории Южного федерального университета – специальность, научное направление, дата защиты диссертации, организация, на базе которой проходила защита.

Метаописания авторефератов электронного архива Новосибирского государственного университета выглядят следующим образом:

nsu-diss.degree к. ф.-м. н. ru_RU

nsu-diss.opponent Никитин Николай Васильевич, д. ф.-м. н. ru_RU

nsu-diss.opponent Шторк Сергей Иванович, к. ф.-м. н. ru_RU

nsu-diss.leading-organisation Центральный аэрогидродинамический институт имени профессора Н. Е. Жуковского ru_RU

nsu-diss.head Катасонов Михаил Михайлович, д. ф.-м. н. ru_RU

nsu-diss.speciality.code 01.02.05 – механика жидкости, газа и плазмы ru_RU

nsu-diss.workorganisation Институт теоретической и прикладной механики им. С. А. Христиановича ru_RU

В метаданных на ВКР цифрового репозитория Южного федерального университета указаны специальность/направление подготовки, год защиты, квалификация выпускника, научный руководитель, институт/академия, факультет, кафедра; в метаданных на студенческие работы – образовательная программа, учебный год, дисциплина.

В метаописаниях на патенты электронного научного архива Уральского федерального университета и метаданных на наборы данных цифрового репозитория Южного федерального университета названия, аннотации и ключевые слова продублированы на английском языке. В репозитории Башкирского государственного медицинского университета все метаданные дублируются на английском языке.

Эта информация полезна для понимания на зарубежных платформах при интеграции репозиториев в международные агрегаторы.

Однако изучение метаданных ИР показало, что, как правило, они не в достаточной степени отражают специфическую информацию на нетекстовые объекты. Например, Электронная библиотечная система

Алтайского государственного университета среди различных объектов интеллектуальной собственности включает базы данных (БД). В метаданных на них не представлена информация о том, в какой программе ведётся БД. Цифровой репозиторий Южного федерального университета содержит наборы данных, в метаданных которых отражена только общая информация. Отсутствуют характеристика содержания этих наборов, тип содержащихся файлов, вид и объём данных, системные требования, наличие цвета и звука и т. п. (как это предусмотрено при составлении библиографических записей в электронных каталогах).

Изученные репозитории образовательных организаций не выкладывают аудио- и видеоконтент в ОД. Однако ИР других учреждений минимальный набор метаданных на данный контент предоставляют. Например, характеристика видеозаписей семинаров и аудиозаписей докладов и эфиров на радио «Каталог публикаций» Института народнохозяйственного прогнозирования ограничивается сведениями об авторе, названии и краткой аннотацией. Отсутствует информация о длительности, формате аудио- и видеозаписей, тематике, целевом назначении, дате и месте проведения мероприятия.

Анализ метаданных объектов изученных ИР позволил сделать следующие выводы:

метаданные, как правило, включают базовые элементы с описательной информацией: автор, заглавие, тип документа, язык, даты. Некоторые метаданные включают сведения только о первом авторе, не отражая соавторов, редакторов и сведения о коллективах;

38% репозиториев в метаданных не предоставляют сведения об источнике (из-за чего при поиске информации по источнику или соавтору пользователь не получит в результатах поиска информацию об этих документах);

структурные данные отражены не во всех метаданных объектов ИР;

за исключением редких случаев не отражены административные элементы, включающие сведения о правах доступа, почти нигде не отражены элементы relation/связь, coverage/охват;

в большинстве метаданных ИР заполнен элемент subject, однако, как правило, в нём даны только ключевые слова, классификационные индексы предоставлены лишь в нескольких репозиториях.

Влияние системы метаданных на эффективность использования репозиториев образовательных организаций

В пленарном докладе на LIBCOM-2023 Я. Л. Шрайберг отметил, что в университетских библиотеках должен быть высокоуровневый и эффективный доступ к научным публикациям и образовательному пространству [**26**. С. 8]. Большую роль в структуре открытой науки играют ИР, они являются одним из популярных инструментов для обеспечения ОД к публикациям. Однако для эффективного использования контента ИР необходимо обеспечить не только доступность информации, но и видимость контента, интуитивно понятный способ отслеживания и поиска различных типов документов. Для этого необходимо предоставлять как можно больше описательной, структурной, административной информации о контенте: о всех лицах, участвующих в создании документа, об источниках публикации – для возможности поиска по этим элементам; о правах доступа – для определения возможности доступа к контенту и быстрого получения полного текста документа, в случае его доступности; давать расширенные аннотации/рефераты – для быстрого определения содержания документа, его релевантности; включать классификационные индексы для повышения эффективности тематического поиска.

Структуру метаданных также необходимо дополнить специфическими сведениями для нетекстовых объектов репозиториев. В частности, необходимо предоставлять информацию о длительности, формате аудио- и видеозаписей, тематике, целевом назначении, дате и месте проведения мероприятий; указывать типы содержащихся файлов в наборах данных, обозначать их вид и объём, системные требования, наличие цвета и звука и т. п.

Дублирование основных элементов метаданных на английском языке будет способствовать лучшему восприятию информации при предоставлении этих сведений в международных агрегаторах, ускорению процесса международного доступа к контенту.

Составление метаданных является достаточно квалифицированным трудом и требует дополнительных трудовых и временных затрат. Но, поскольку системы метаданных являются одним из важнейших компонентов ИР, значительно влияют на их поисковые возможности, их качество и полнота ввода элементов являются важным направлени-

ем деятельности. Хорошо организованные метаданные позволяют оптимизировать поисковые возможности объектов ИР, организовать полнотекстовый и расширенный поиск информации с применением фасетной навигации, булевых операторов, управлением результатами поиска.

Усложнение структуры метаданных, большее разнообразие и более подробное описание элементов метаданных повысит эффективность информационного и тематического поиска информации, функциональные возможности ИР.

Список источников

1. **Bankier J.-G., Perciali I.** The institutional repository rediscovered: what can a university do for open access publishing? // Serials Review. 2008. Vol. 34, № 1. P. 21–26.

2. **White W.** Institutional repositories: contributing to institutional knowledge management and the global research commons // 4th International Open Repositories Conference : paper presented, 18–21 May, Atlanta. 2009. URL: http://eprints.soton.ac.uk/48552/ (дата обращения: 19.04.2024).

3. **Daly R., Organ M.** Research online: digital commons as a publishing platform at the University of Wollongong // Australia Serials Review. 2009. Vol. 35. № 3. P. 149–153.

4. **Lyte V., Jones S., Ananiadou S., Kerr L**. UK institutional repository search: innovation and discovery. 2009. URL: www.ariadne.ac.uk/issue61/lyte-et-al (дата обращения: 19.04.2024).

5. **Ovadia S.** When social media meets scholarly publishing // Behavioral & Social Sciences Librarian. 2013. Vol. 32, № 3. P. 194–198.

6. **Редькина Н. С.** Современные тенденции и проблемы развития репозиториев // Информационные ресурсы России. 2022. № 2 (186). С. 81–93. https://doi.org/10.52815/0204-3653_2022_021.

7. **Засурский И. И., Соколова Д. В., Трищенко Н. Д.** Репозитории открытого доступа: функции и тенденции развития // Научные и технические библиотеки. 2020. № 9. С. 121–142. URL: https://ntb.gpntb.ru/jour/article/view/652 (дата обращения: 19.04.2024).

8. **Юдина И. Г., Федотова О. А.** Репозитории научных публикаций открытого доступа: история и перспективы развития // Информационное общество. 2020. № 6. С. 67–79. URL: http://infosoc.iis.ru/article/view/514/435 (дата обращения: 19.04.2024).

9. **Шварцман М. Е.** К вопросу выбора программного обеспечения для институционального репозитория // Science Online: электронные информационные ресурсы для науки и образования : 17-я Международная конференция, о. Сицилия, 24–31 мая, 2013. Москва, 2013. С. 42.

10. **Васильева Н. В.** Открытое программное обеспечение для организации репозитория // Научные и технические библиотеки. 2023. № 3. С. 102–119. https://doi.org/10.33186/1027-3689-2023-3-102-119.

11. **Ковязина Е. В.** Репозиторий научных публикаций: проблемы обмена данными // Science Online: электронные информационные ресурсы для науки и образования : 19-я Международная конференция, Андорра, 17–24 января 2015. Москва. 2015. URL: http://elibrary.ru/projects/conference/andorra2015/presentations/2015–01–22/kovyazina.pps (дата обращения: 17.04.2024).

12. **Стукалова А. А.** Функциональные возможности репозиториев вузов – участников программы «Приоритет-2030» // Труды ГПНТБ СО РАН. 2022. № 2. С. 36–47. https://doi.org/10.20913/2618-7515-2022-2-36-47.

13. **Гончаров М. В., Колосов К. А.** Анализ метаданных российских репозиториев открытого доступа по научно-технической тематике с целью их использования в системе единого открытого архива информации ГПНТБ России // Научные и технические библиотеки. 2021. № 12. С. 15–28. URL: https://elibrary.ru/item.asp?id=47557987 (дата обращения: 17.04.2024).

14. **Резниченко О. С.** Подготовка публикации метаданных для пакетного импорта в институциональные репозитории, основанные на DSpace // Экономика. Информатика. 2021. № 48 (3). С. 564–577. DOI 10.52575/2687-0932-2021-48-3-564-577.

15. **Литвинова Н. Н.** Методические рекомендации для разработчиков по организации работы открытых репозиториев: корректное создание метаданных. Москва : Ваше цифровое издательство, 2020. 20 с. URL: https://www.openrepository.ru/images/docs/Metod_Litvinova_v3.pdf (дата обращения: 17.04.2024).

16. **Зуйкина К. Л., Соколова Д. В., Скалабан А. В.** Электронные библиотеки в России. Текущий статус и перспективы развития. Москва : Ваш формат, 2017. 120 с. URL: https://upload.wikimedia.org/wikipedia/commons/5/56/%D0%AD%D0%BB%D0%B5%D0%BA%D1%82%D1%80%D0%BE%D0%BD%D0%BD%D1%8B%D0%B5_%D0%B1%D0%B8%D0%B1%D0%BB%D0%B8%D0%BE%D1%82%D0%B5%D0%BA%D0%B8_%D0%B2_%D0%A0%D0%BE%D1%81%D1%81%D0%B8%D0%B8.pdf (дата обращения: 17.04.2024).

17. **Мазов Н. А., Жижимов О. Л.** Метаданные и их роль в распределённых информационных системах на основе использования протокола Z39.50. Лекция // Библиосфера. 2006. № 2. С. 51–59. URL: https://cyberleninka.ru/article/n/metadannye-i-ih-rol-v-raspredelennyh-informatsionnyh-sistemah-na-osnove-ispolzovaniya-protokola-z39-50-lektsiya (дата обращения: 17.04.2024).

18. **Антопольский А. Б.** Системы метаданных в электронных библиотеках // Научные и технические библиотеки. 2002. № 3. С. 26–44. URL: https://www.gpntb.ru/win/ntb/ntb2002/3/f3_04.htm (дата обращения: 17.04.2024).

19. **ГОСТ Р 55750-2013.** Метаданные электронных образовательных ресурсов: общие положения. Москва : Стандартинформ, 2014. 6 с. URL: https://internet-law.ru/gosts/gost/55798/ (дата обращения: 17.04.2024).

20. **Берестова Т. Ф.** Что скрывается за термином «Метаданные»? // Вестник ЧГАКИ. 2017. № 1 (49). URL: https://cyberleninka.ru/article/n/chto-skryvaetsya-za-terminom-metadannye (дата обращения: 17.04.2024).

21. **Негуляев Е. А.** Универсальные схемы метаданных и задача описания веб-ресурсов // Библиосфера. 2006. № 1. С. 43–54. URL: https://cyberleninka.ru/article/n/universalnye-shemy-metadannyh-i-zadacha-opisaniya-veb-resursov (дата обращения: 17.04.2024).

22. **Деденок Ю. Л.** Создание электронной библиотеки научных трудов на платформе DSpace. URL: https://vdocuments.mx/-5875f0b01a28ab006e8b4b7b.html?page=1 (дата обращения: 18.04.2024).

23. **ГОСТ Р 7.0.10-2019.** Система стандартов по информации, библиотечному и издательскому делу. Набор элементов метаданных «Дублинское ядро»: основные (ядерные) элементы. Москва : Стандартинформ, 2019. 7 с. URL: https://files.stroyinf.ru/Data/707/70766.pdf (дата обращения: 17.04.2024).

24. **Литвинова Н. Н.** Метаданные репозитория организации: проблемы интеграции с внешними системами. URL: https://conf.neicon.ru/materials/77-online0520/20200528-Litvinova.pdf (дата обращения: 17.04.2024).

25. **Прокудин Д. Е., Чугунов А. В.** Сетевые электронные репозитории, системы агрегации и сохранения в концепции свободного доступа к научной информации // Интегрированные цифровые ресурсы: организационно-технологические и научно-методические основы развития. 2015. Вып. 6. 271 с. URL: https://dspace.spbu.ru/bitstream/11701/36449/1/%d0%9f%d1%80%d0%be%d0%ba%d1%83%d0%b4%d0%b8%d0%bd%d0%94%d0%95_%d0%a7%d1%83%d0%b3%d1%83%d0%bd%d0%be%d0%b2%d0%90%d0%92_%d0%9f%d1%80%d0%b5%d0%b7%d0%b8%d0%b4%d0%b5%d0%bd%d1%82%d1%81%d0%ba%d0%b0%d1%8f_%d0%b1%d0%b8%d0%b1%d0%bb%d0%b8%d0%be%d1%82%d0%b5%d0%ba%d0%b0_2015.pdf (дата обращения: 17.04.2024).

26. **Шрайберг Я. Л.** Цифровая трансформация библиотек и образовательных учреждений: основные компоненты и вызовы: пленарный доклад. Москва : ГПНТБ России, 2023. 24 с.

References

1. **Bankier J.-G., Perciali I.** The institutional repository rediscovered: what can a university do for open access publishing? // Serials Review. 2008. Vol. 34, № 1. P. 21–26.

2. **White W.** Institutional repositories: contributing to institutional knowledge management and the global research commons // 4th International Open Repositories Conference : paper presented, 18–21 May, Atlanta. 2009. URL: http://eprints.soton.ac.uk/48552/ (дата обращения: 19.04.2024).

3. **Daly R., Organ M.** Research online: digital commons as a publishing platform at the University of Wollongong // Australia Serials Review. 2009. Vol. 35. № 3. P. 149–153.

4. **Lyte V., Jones S., Ananiadou S., Kerr L**. UK institutional repository search: innovation and discovery. 2009. URL: www.ariadne.ac.uk/issue61/lyte-et-al (дата обращения: 19.04.2024).

5. **Ovadia S.** When social media meets scholarly publishing // Behavioral & Social Sciences Librarian. 2013. Vol. 32, № 3. P. 194–198.

6. **Red`kina N. S.** Sovremenny`e tendentcii i problemy` razvitiia repozitoriev // Informatcionny`e resursy` Rossii. 2022. № 2 (186). S. 81–93. https://doi.org/10.52815/0204-3653_2022_021.

7. **Zasurskii` I. I., Sokolova D. V., Trishchenko N. D.** Repozitorii otkry`togo dostupa: funktcii i tendentcii razvitiia // Nauchny`e i tekhnicheskie biblioteki. 2020. № 9. S. 121–142. URL: https://ntb.gpntb.ru/jour/article/view/652 (data obrashcheniia: 19.04.2024).

8. **Iudina I. G., Fedotova O. A.** Repozitorii nauchny`kh publikatcii` otkry`togo dostupa: istoriia i perspektivy` razvitiia // Informatcionnoe obshchestvo. 2020. № 6. S. 67–79. URL: http://infosoc.iis.ru/article/view/514/435 (data obrashcheniia: 19.04.2024).

9. **Shvartcman M. E.** K voprosu vy`bora programmnogo obespecheniia dlia institutcional`nogo repozitoriia // Science Online: e`lektronny`e informatcionny`e resursy` dlia nauki i obrazovaniia : 17-ia Mezhdunarodnaia konferentciia, o. Sitciliia, 24–31 maia, 2013. Moskva, 2013. S. 42.

10. **Vasil`eva N. V.** Otkry`toe programmnoe obespechenie dlia organizatcii repozitoriia // Nauchny`e i tekhnicheskie biblioteki. 2023. № 3. S. 102–119. https://doi.org/10.33186/1027-3689-2023-3-102-119.

11. **Koviazina E. V.** Repozitorii` nauchny`kh publikatcii`: problemy` obmena danny`mi // Science Online: e`lektronny`e informatcionny`e resursy` dlia nauki i obrazovaniia : 19-ia Mezhdunarodnaia konferentciia, Andorra, 17–24 ianvaria 2015. Moskva. 2015. URL: http://elibrary.ru/projects/conference/andorra2015/presentations/2015–01–22/kovyazina.pps (data obrashcheniia: 17.04.2024).

12. **Stukalova A. A.** Funktcional`ny`e vozmozhnosti repozitoriev vuzov – uchastneykov programmy` «Prioritet-2030» // Trudy` GPNTB SO RAN. 2022. № 2. S. 36–47. https://doi.org/10.20913/2618-7515-2022-2-36-47.

13. **Goncharov M. V., Kolosov K. A.** Analiz metadanny`kh rossii`skikh repozitoriev otkry`togo dostupa po nauchno-tekhnicheskoi` tematike s tcel`iu ikh ispol`zovaniia v sisteme edinogo otkry`togo arhiva informatcii GPNTB Rossii // Nauchny`e i tekhnicheskie biblioteki. 2021. № 12. S. 15–28. URL: https://elibrary.ru/item.asp?id=47557987 (data obrashcheniia: 17.04.2024).

14. **Reznichenko O. S.** Podgotovka publikatcii metadanny`kh dlia paketnogo importa v institutcional`ny`e repozitorii, osnovanny`e na DSpace // E`konomika. Informatika. 2021. № 48 (3). S. 564–577. DOI 10.52575/2687-0932-2021-48-3-564-577.

15. **Leetvinova N. N.** Metodicheskie rekomendatcii dlia razrabotchikov po organizatcii raboty` otkry`ty`kh repozitoriev: korrektnoe sozdanie metadanny`kh. Moskva : Vashe tcifrovoe izdatel`stvo, 2020. 20 s. URL: https://www.openrepository.ru/images/docs/Metod_Litvinova_v3.pdf (data obrashcheniia: 17.04.2024).

16. **Zui`kina K. L., Sokolova D. V., Skalaban A. V.** E`lektronny`e biblioteki v Rossii. Tekushchii` status i perspektivy` razvitiia. Moskva : Vash format, 2017. 120 s. URL: https://upload.wikimedia.org/wikipedia/commons/5/56/%D0%AD%D0%BB%D0%B5%D0%BA%D1%82%D1%80%D0%BE%D0%BD%D0%BD%D1%8B%D0%B5_%D0%B1%D0%B8%D0%B1%D0%BB%D0%B8%D0%BE%D1%82%D0%B5%D0%BA%D0%B8_%D0%B2_%D0%A0%D0%BE%D1%81%D1%81%D0%B8%D0%B8.pdf (data obrashcheniia: 17.04.2024).

17. **Mazov N. A., Zhizhimov O. L.** Metadanny`e i ikh rol` v raspredelyonny`kh informatcionny`kh sistemakh na osnove ispol`zovaniia protokola Z39.50. Lektciia // Bibliosfera. 2006. № 2. S. 51–59. URL: https://cyberleninka.ru/article/n/metadannye-i-ih-rol-v-raspredelennyh-informatsionnyh-sistemah-na-osnove-ispolzovaniya-protokola-z39-50-lektsiya (data obrashcheniia: 17.04.2024).

18. **Antopol`skii` A. B.** Sistemy` metadanny`kh v e`lektronny`kh bibliotekakh // Nauchny`e i tekhnicheskie biblioteki. 2002. № 3. S. 26–44. URL: https://www.gpntb.ru/win/ntb/ntb2002/3/f3_04.htm (data obrashcheniia: 17.04.2024).

19. **GOST R 55750-2013.** Metadanny`e e`lektronny`kh obrazovatel`ny`kh resursov: obshchie polozheniia. Moskva : Standartinfom, 2014. 6 s. URL: https://internet-law.ru/gosts/gost/55798/ (data obrashcheniia: 17.04.2024).

20. **Berestova T. F.** Chto skry`vaetsia za terminom «Metadanny`e»? // Vestneyk ChGAKI. 2017. № 1 (49). URL: https://cyberleninka.ru/article/n/chto-skryvaetsya-za-terminom-metadannye (data obrashcheniia: 17.04.2024).

21. **Neguliaev E. A.** Universal`ny`e skhemy` metadanny`kh i zadacha opisaniia veb-resursov // Bibliosfera. 2006. № 1. S. 43–54. URL: https://cyberleninka.ru/article/n/universalnye-shemy-metadannyh-i-zadacha-opisaniya-veb-resursov (data obrashcheniia: 17.04.2024).

22. **Dedenok Iu. L.** Sozdanie e`lektronnoi` biblioteki nauchny`kh trudov na platforme DSpace. URL: https://vdocuments.mx/-5875f0b01a28ab006e8b4b7b.html?page=1 (data obrashcheniia: 18.04.2024).

23. **GOST R 7.0.10-2019.** Sistema standartov po informatcii, bibliotechnomu i izdatel`skomu delu. Nabor e`lementov metadanny`kh «Dublinskoe iadro»: osnovny`e (iaderny`e) e`lementy`. Moskva : Standartinform, 2019. 7 s. URL: https://files.stroyinf.ru/Data/707/70766.pdf (data obrashcheniia: 17.04.2024).

24. **Leetvinova N. N.** Metadanny`e repozitoriia organizatcii: problemy` integratcii s vneshnimi sistemami. URL: https://conf.neicon.ru/materials/77-online0520/20200528-Litvinova.pdf (data obrashcheniia: 17.04.2024).

25. **Prokudin D. E., Chugunov A. V.** Setevy`e e`lektronny`e repozitorii, sistemy` agregatcii i sokhraneniia v kontceptcii svobodnogo dostupa k nauchnoi` informatcii // Integrirovanny`e tcifrovy`e resursy`: organizatcionno-tekhnologicheskie i nauchno-metodicheskie osnovy` razvitiia. 2015. Vy`p. 6. 271 s. URL: https://dspace.spbu.ru/bitstream/11701/36449/1/%d0%9f%d1%80%d0%be%d0%ba%d1%83%d0%b4%d0%b8%d0%bd%d0%94%d0%95_%d0%a7%d1%83%d0%b3%d1%83%d0%bd%d0%be%d0%b2%d0%90%d0%92_%d0%9f%d1%80%d0%b5%d0%b7%d0%b8%d0%b4%d0%b5%d0%bd%d1%82%d1%81%d0%ba%d0%b0%d1%8f_%d0%b1%d0%b8%d0%b1%d0%bb%d0%b8%d0%be%d1%82%d0%b5%d0%ba%d0%b0_2015.pdf (data obrashcheniia: 17.04.2024).

26. **Shrai`berg Ia. L.** Tcifrovaia transformatciia bibliotek i obrazovatel`ny`kh uchrezhdenii`:osnovny`e komponenty` i vy`zovy`: plenarny`i` doclad. Moskva : GPNTB Rossii, 2023. 24 s.

Информация об авторе / Author

Стукалова Анна Александровна – канд. пед. наук, старший научный сотрудник ГПНТБ СО РАН, Новосибирск, Российская Федерация

Stukalova@gpntbsib.ru

Anna A. Stukalova – Cand. Sc. (Pedagogy), Senior Researcher, State Public Scientific and Technological Library of the Russian Academy of Sciences Siberian Branch, Novosibirsk, Russian Federation

Stukalova@gpntbsib.ru