

ОТКРЫТАЯ
НАУКА
РОССИИ

Метаданные репозитория организации: проблемы интеграции с внешними системами



Наталия Николаевна Литвинова
Nlit@neicon.ru

Что такое открытый репозиторий?

Открытые репозитории получили распространение в рамках движения Открытого доступа к научной литературе. Размещение работ в открытом репозитории, или самоархивирование, - классическая модель Зеленого Открытого доступа.

Определение OpenAIRE – европейского проекта поддержки открытого доступа:



“An Open Access repository is a database or a virtual archive established to collect, disseminate and preserve scientific output like scientific articles and datasets and make them freely available”.

Репозиторий открытого доступа – база данных или виртуальный архив, которые создаются для сбора, распространения, сохранения и предоставления в свободный доступ результатов научной деятельности: научных статей и наборов данных.

Репозиторий организации – это репозиторий, ориентированный на научную продукцию данной организации. Другой распространенный тип репозитория – тематические (предметные), в которых размещаются работы сходной тематики, подготовленные в разных организациях.

Репозитории-агрегаторы – репозитории, собирающие метаданные (иногда – и полные тексты) из нескольких репозиториях.

Возможности интеграции репозиториев в рамках проекта «Открытая наука России»

НОРА – Национальный агрегатор открытых репозиториев

<https://www.openrepository.ru/>

22 репозитория России и Белоруссии; около 400 тысяч записей; содержит только метаданные: обращение к полным текстам – из репозиториев-участников.

BASE - Bielefeld Academic Search Engine <https://www.base-search.net/>

Более 7 тысяч участников; более 150 миллионов записей (с дублями); содержит только метаданные: обращение к полным текстам – из репозиториев-участников. НОРА – участник BASE.

OpenAIRE – European Open Science Infrastructure (сейчас – глобальный проект)

<https://explore.openaire.eu/>

Около 17 тысяч участников; более 30 миллионов записей (дубли сводятся к единой записи); содержит метаданные и частично – полные тексты **только для целей внутрисистемного анализа**. НОРА – участник OpenAIRE

CORE – Connecting Repositories <https://core.ac.uk/>

Около 10 тысяч участников; около 180 миллионов записей; содержит метаданные и полные тексты. **НОРА не поставляет данные в CORE, команда регистрирует репозитории-участники.**

Что дает интеграция репозиториям-участникам?

1. Увеличение видимости объектов репозитория
 - 1.1. Увеличение количества обращений к ресурсам из агрегаторов.
 - 1.2. Индексирование агрегаторов в глобальных сервисах дискавери способствует появлению ресурсов репозитория в поисковой выдаче сотен библиотек по всему миру. Контент BASE индексируется сервисом дискавери EBSCO Discovery Service, Контент CORE – дискавери-сервисами Primo и Summon.

Иллюстрация: поиск в дискавери-сервисе EDS в РГБ по теме «лазерная ангиопластика»

Поиск: Discovery Service - Российская государственная библиотека

лазерная ангиопластика

AND

AND

Базовый поиск Расширенный поиск История поиска

Уточнить результаты

Текущий поиск

Найти все термины: лазерная ангиопластика

Ограничение до

Полный текст (fulltext)

Научные (экспертные) журналы

Известно в библиотечной коллекции

2007 2019

Полная библиография

Авторы

Источники

Все результаты

Научные журналы (3)

Электронные ресурсы (3)

Обзоры (1)

Полная библиография

Тема

Издание

Результаты поиска: 1 - 17 из 17

Значимость Параметры страницы Поделись

Примечание: Из результатов поиска удалены патенты

1. КВАНТОВО-КАСКАДНЫЕ ЛАЗЕРЫ В МЕДИЦИНЕ И БИОЛОГИИ (Обзор) / QUANTUM-CASCADE LASERS IN MEDICINE AND BIOLOGY (Review)

By: Абрам, П.И. / Abram, P.I.; Кузнец, Е.В. / Kuznetsov, E.V.; Саваров, Л.А. / Savarov, L.A.; Саваров, М.И. / Savarov, M.I. In: Журнал прикладной спектроскопии. 2019. (1):5-32. Государственное научное учреждение Институт физики им. Б.Н. Степанова Национальной академии наук Беларуси. Language: Russian. Без данных: ELibRARY.RU

Обзор

Рассмотрены вопросы применения квантово-каскадных лазеров (ККЛ) в медицинской практике. Проведен сравнительный анализ методов лазерной спектроскопии для неинвазивной диагностики заболеваний.

Тематика: квантово-каскадные лазеры, квантовая биология, инфракрасная спектроскопия, оптоэлектроника, анализ данных, лазерный скайлер, лазерная ангиопластика, транскатодный излучатель, микроэлектроника, инфракрасная спектроскопия, инфракрасная спектроскопия, pump-probe technology, Quantum cascade lasers, Soft biological tissue, Infrared absorption spectroscopy, breath analysis, laser scalpel, laser angioplasty, lenselet imaging, Microspectroscopy, infrared hyperspectral imaging, pump-probe technology

Просмотреть название полного текста на eLibrary

2. Транскатодный лазерный ревазкуляризация миокарда как альтернативный метод лечения при ишемической болезни сердца

Эксперименты By: Г.К. АУТЮКИН. In: Вестник Казахского Национального медицинского университета. 1995-01-30:4-3. 19-75. Казахстан. Алматы. Республиканское государственное предприятие на праве хозяйственного ведения «АКНМД» им. С.Д. Асфендиярова. Министерство здравоохранения и социального развития Республики Казахстан. 2015. Language: Russian. Без данных: CyberLeninka

Аннотация: Транскатодный лазерный ревазкуляризация миокарда (ТЛР) является новым стандартом в хирургическом лечении ишемической болезни сердца. Но у больных с диффузными изменениями в коронарных сосудах или коронарии и пр.

Тематика: ТМР-транскатодный лазерный ревазкуляризация; ТЛР - транскатодный ангиопластика

3. Использование лазеров в сердечно-сосудистой хирургии: от эксперимента к практике

By: Шутлин, Владимир Митрофанович; Андреев, С. Л.; Вечерский, Юрий Юрьевич; Гордов, Евгений Петрович. STT Language: Russian. Без данных: BASE

Тематика: лазерная ангиопластика, транскатодный лазерный ревазкуляризация, полупроводниковые лазеры

Просмотреть название полного текста в открытой базе BASE

ЭЛЕКТРОННАЯ БИБЛИОТЕКА (РЕПОЗИТОРИЙ)
Томского государственного университета

Home Show All 65975 Show Quick Collection 0

Бrowse Search History

Эндаумент фонд ТГУ! Формируя фонд в истории университета - Ссылка пожертвования

Search

Advanced Search

Add to Quick Collection

Description	Size	Format
Использование лазеров в сердечно-сосудистой хирургии: от эксперимента к практике	4 MB	Adobe Acrobat PDF

Title: Использование лазеров в сердечно-сосудистой хирургии: от эксперимента к практике

Creator: Шутлин, Владимир Митрофанович

Creator: Андреев, С. Л.

Creator: Вечерский, Юрий Юрьевич

Creator: Гордов, Евгений Петрович

Subject

Date: 2010

Publisher: Томск : STT

Relationships

collection(s): Физический факультет

Identifier

Rights: Resource access is available in the premises of Research Library

Type: монография

Language: rus

Preview

Использование лазеров в сердечно-сосудистой хирургии: от эксперимента к практике / Шутлин В. М., Андреев С. Л., Вечерский Ю. Ю., Гордов Е. П. и др. - Томск : STT, 2010. URL: <http://vital.lib.tsu.ru/vital/access/manager/Repository/VTLS000437066>

Статистика перехода по ссылкам в РГБ за 2019 год: всего – 24113, из них на BASE – 6404 (27%)

Что дает интеграция репозиториям-участникам?

2. Интеграция с другими глобальными сервисами

Контент агрегаторов учитывается сервисами Unpaywall, Open Access Button и другими сервисами определения свободного доступа к контенту.

BASE интегрирован с ORCID: автор публикации может воспользоваться опцией добавления ее в список своих работ в ORCID.

CORE позволяет внедрить на страницы репозитория CORE Recommender, который выводит список похожих документов для любого найденного в репозитории.

МЕДИАОБРАЗОВАТЕЛЬНЫЙ КОНТЕКСТ УЛИЧНОГО ИСКУССТВА КАК ФОРМЫ ВИЗУАЛЬНОЙ КОММУНИКАЦИИ

Онкович Анна Владимировна, Онкович Артем Дмитриевич

[ПРОСМОТР](#)

Дата публикации: 2016-01-01

Дата публикации в реестре: 2020-02-29T23:47:34Z

Аннотация:

Среди разновидностей визуальной коммуникации ныне бурно развивается уличное искусство - стрит-арт (англ. street art), характерной особенностью которого является ярко выраженный урбанистический стиль с включением множества художественных методик и материалов. В докладе делается попытка рассмотреть относительно новые явления - стрит-арти мурал-арт - с позиций медиаобразования для осознания возможностей стрит-арт-дидактики и муралдидактики в учебном процессе. Уличное визуальное искусство как средство массовой коммуникации приобретает различные формы, поэтому муралдидактике как части стрит-арт-дидактики должно принадлежать особое место в образовательном процессе.

Ключевые слова:

визуальная коммуникация, уличное искусство, изобразительное искусство; медиаобразование, медиаобразовательный аспект; стрит-арт-дидактика, муралдидактика, visual communication, street art, visual arts; media education, media education aspect; Street?art?didactics, mural didactics

Тип: article

Связанные документы (рекомендация CORE)

- 

Street art as an object of studying in various sociological concepts
 Provided by: Institutional repository of Ural Federal University named after the first President of Russia B.N.Yeltsin | Publisher: Уральский федеральный университет | Year: 2018
 by Shvindt U., Швиндт У. С.
- 

Организация самостоятельной работы обучающихся в процессе изучения явлений современного искусства
 Provided by: Electronic archive USPU | Year: 2018
 by Чернавских О.С.
- 

Елементи стріт-арту у семантико-синтагматичному просторі міста
 Provided by: Scientific-theoretical almanac "Grani" / Науково-теоретичний альманах "Грані" | Publisher: Publishing house «Grani» | Year: 2019
 by Міхно Н. К.

Для описания ресурсов репозиториев используется единый формат метаданных Dublin Core (DC) – Дублинское ядро, разработанный в рамках движения Open Access

Краткая история

1995 год – семинар в Дублине, Огайо; обсуждение семантики метаданных для описания научных публикаций в вебе.

1996 год – публикация набора метаданных из **15 элементов** Dublin Core Metadata Element Set (DCMES)

2001 год – схема метаданных опубликована как национальный стандарт США NISO Z39.85-2001

2003 год – принят международный стандарт - ISO 15836-2003

2017 год – последняя редакция - ISO 15836-1:2017

2019 год - ГОСТ Р 7.0.10-2019 (ИСО 15836-1:2017) – переводная адаптированная версия с комментариями, учитывающими особенности российской национальной стандартизации.

15 элементов Дублинского ядра составляют так называемый простой (базовый) набор метаданных. Именно он является основой для обмена информацией между репозиториями и интеграции их контента в агрегаторах.

Расширенный (квалифицированный) набор метаданных содержит 50+ термов, включающих 15 основных элементов и термы, детализирующие семантику базового набора. Стандарт ISO 15836-2:2019

Базовый набор элементов DC

(русский эквивалент и толкование – по ГОСТ Р 7.0.10-2019)

Название	Русский эквивалент	Толкование
Contributor	<i>Соисполнитель</i>	Лицо, ответственное за вклад в создание ресурса
Coverage	<i>Охват</i>	Характеристика содержания ресурса в пространстве или времени
Creator	<i>Создатель</i>	Лицо, несущее первичную ответственность за создание ресурса
Date	<i>Дата</i>	Момент или отрезок времени, связанный с событием жизненного цикла ресурса
Description	<i>Описание</i>	Сведения о ресурсе
Format	<i>Формат</i>	Формат файла, физический носитель или размеры ресурса
Identifier	<i>Идентификатор</i>	Неразрывная строка символов, идентифицирующая ресурс
Language	<i>Язык</i>	Язык содержательной части ресурса
Publisher	<i>Издатель</i>	Лицо, ответственное за предоставление ресурса пользователям
Relation	<i>Связь</i>	Связанный (родственный) ресурс
Rights	<i>Права</i>	Сведения о правах на использование и управление ресурсом
Source	<i>Источник</i>	Связанный (родственный) ресурс, послуживший источником данных для описываемого ресурса
Subject	<i>Предмет</i>	Тематика ресурса, предметное содержание ресурса
Title	<i>Заглавие</i>	Имя, данное ресурсу
Type	<i>Тип ресурса</i>	Природа или жанр содержания ресурса

Проблемы семантики базового набора DC.

Возможности детализации

Важное достоинство базового набора элементов DC – в том, что его широкая семантика дает возможность описания любых веб-ресурсов.

Оборотные стороны этого достоинства:

1. Широта семантики дает простор для различных интерпретаций. Как отмечается в литературе, наиболее сложными в интерпретации и разграничении являются пары “relation – source”, “creator – contributor”, “type – format”.
2. Для описания конкретных категорий ресурсов глобальный уровень оказывается недостаточным: он не отражает важные характеристики ресурса. Это в полной мере относится к основным объектам описания в репозиториях – статьям, материалам конференций, книгам, диссертациям.

Более детальные элементы описания ресурсов могут вводиться с использованием:

- расширенного набора термов DC, включенных в стандарт ISO 15836 Part 2: "DCMI Properties and classes" (ISO 15836-2:2019);
- других форматов метаданных, таких как MODS (Metadata Object Description Schema) на базе упрощенного набора элементов формата MARC, ETD-MS для описания диссертаций, DataCite Metadata Schema и других;
- собственных наборов метаданных, как правило, формируемых на основе распространенного формата с добавлением нескольких специфических элементов.

Сбор метаданных по протоколу OAI-PMN для обмена информацией

Сбор метаданных для ресурс-агрегатора удобно организовать с использованием одного выбранного формата. Например, если агрегатор собирает метаданные только для диссертаций, участники могут договориться применять единый формат описания диссертаций, который будет использоваться для сбора; другие форматы игнорируются.

Если ресурс-агрегатор не использует специализированный формат метаданных, разделяемый всеми участниками, то универсальным форматом для сбора становится базовый набор DC (префикс `oai_dc`). Этот набор используется также в случаях, когда предложенный агрегатором специализированный формат не поддерживается отдельными репозиториями - поставщиками данных. Такая ситуация характерна, например, для агрегатора OpenAIRE: если метаданные не представлены в специальном сете (с 4-й версии – префикс `oai_openaire`), то они собираются из базового набора DC.

При сборе метаданных в НОРА и передаче их в международные агрегаторы используется базовый набор DC, префикс `oai_dc`.

При этом происходит некоторое обобщение данных репозиторий-поставщиков.

Сравнение метаданных в репозитории с представлением их в обменном формате oai_dc

 Казанский федеральный университет		Электронный архив
Главная → Институты и другие подразделения → Институт математики и механики им. Н.И. Лобачевского Просмотр элемента		
Показать сокращенную информацию		
dc.contributor	Казанский федеральный университет	
dc.contributor.author	Абызов Аделъ Наилевич	
dc.contributor.author	Арсланов Марат Мирзаевич	
dc.date.accessioned	2019-04-05T07:06:14Z	
dc.date.available	2019-04-05T07:06:14Z	
dc.date.issued	2019	
dc.identifier.citation	А.Н. Абызов. Исследования по алгебре и математической логике в Казанском университете / А.Н. Абызов, М.М. Арсланов // Труды семинара кафедры алгебры и математической логики Казанского (Приволжского) федерального университета, Итоги науки и техн. Сер. Соврем. мат. и ее прил. Темат. обз. - 157. - М.: ВИНТИ РАН, 2018. - С. 3-7.	
dc.identifier.uri	https://dspace.kpfu.ru/xmlui/handle/net/150436	
dc.description.abstract	Кратко представлена история Казанской алгебраической школы.	
dc.language.iso	ru	
dc.relation.ispartofseries	Итоги науки и техники. Современная математика и ее приложения. Тематические обзоры	
dc.rights	открытый доступ	
dc.subject	алгебра	
dc.subject	математическая логика	
dc.subject.other	Математика	
dc.title	Исследования по алгебре и математической логике в Казанском университете	
dc.type	Article	
dc.contributor.org	Институт математики и механики им.Н.И.Лобачевского	
dc.description.pages	3-7	
dc.relation.ispartofseries-volume	157	

```
<dc:title>Исследования по алгебре и математической логике в Казанском университете</dc:title>
```

```
<dc:creator>Абызов Аделъ Наилевич</dc:creator>
```

```
<dc:creator>Арсланов Марат Мирзаевич</dc:creator>
```

```
<dc:contributor>Институт математики и механики им.Н.И.Лобачевского</dc:contributor>
```

```
<dc:contributor>Казанский федеральный университет</dc:contributor>
```

```
<dc:subject>алгебра</dc:subject>
```

```
<dc:subject>математическая логика</dc:subject>
```

```
<dc:subject>Математика</dc:subject>
```

```
<dc:description>Кратко представлена история Казанской алгебраической школы.</dc:description>
```

```
<dc:description>3-7</dc:description>
```

```
<dc:date>2019-04-05T07:06:14Z</dc:date>
```

```
<dc:date>2019-04-05T07:06:14Z</dc:date>
```

```
<dc:date>2019</dc:date>
```

```
<dc:type>Article</dc:type>
```

```
<dc:identifier>А.Н. Абызов. Исследования по алгебре и математической логике в Казанском университете / А.Н. Абызов, М.М. Арсланов // Труды семинара кафедры алгебры и математической логики Казанского (Приволжского) федерального университета, Итоги науки и техн. Сер. Соврем. мат. и ее прил. Темат. обз. - 157. - М.: ВИНТИ РАН, 2018. - С. 3-7.</dc:identifier>
```

```
<dc:identifier>https://dspace.kpfu.ru/xmlui/handle/net/150436</dc:identifier>
```

```
<dc:language>ru</dc:language>
```

```
<dc:relation>Итоги науки и техники. Современная математика и ее приложения. Тематические обзоры</dc:relation>
```

```
<dc:relation>157</dc:relation>
```

```
<dc:rights>открытый доступ</dc:rights>
```

Для обеспечения унификации значений и нужного уровня детализации метаданных, получаемых по OAI-PMH в формате базового DC, репозитории-агрегаторы применяют ряд рекомендаций:

- рекомендации по обязательному использованию некоторых полей;
- унификация использования полей (например, для статей рекомендуется записывать название журнала в поле dc:source);
- унификация формулировок значений некоторых полей, важных для поиска в базе (например, названий типов объектов репозитория, версий документа и т.п.);
- рекомендации по заполнению некоторых полей с заложенными возможностями структурирования информации для последующего разбора и загрузки в отдельные поля базы данных

Примеры

1. OpenAIRE рекомендует в поле dc:type применять следующие значения:

- info:eu-repo/semantics/draft
- info:eu-repo/semantics/submittedVersion
- info:eu-repo/semantics/acceptedVersion
- info:eu-repo/semantics/publishedVersion
- info:eu-repo/semantics/updatedVersion

2. Рекомендации BASE для информации об ORCID автора:

<dc:creator>Robinson, Neil(<https://orcid.org/0000-0002-0893-2190>)</dc:creator>



Учитывая ограничения базового набора DC и специфику сбора метаданных международными агрегаторами BASE и OpenAIRE, команда проекта «Открытая наука России» разработала методические рекомендации по подготовке метаданных для репозиторий, заинтересованных в интеграции во внешние системы.

Документ называется «Методические рекомендации для разработчиков ОР по организации работы репозиторий: корректное создание метаданных».

Он опубликован на сайте проекта в разделе «Методические материалы»

(https://www.openrepository.ru/images/docs/Metod_Litvinova_v3.pdf)

В методике даются рекомендации по наполнению полей базового набора DC и двух дополнительных полей **для сбора их по протоколу OAI-PMH.**

Предлагается сделать обязательными 6 полей (перечисляются с метками формата oai_dc)

dc:rights	права, правовой статус
dc:type	тип, вид, жанр ресурса
dc:creator	создатель, автор
dc:identifier	идентификатор ресурса
dc: title	заглавие
dc:date	дата (создания ресурса)

Обязательность в данном контексте не означает, что записи, в которых этих полей нет, будут отвергнуты при сборе метаданных в НОРА. Мы просто подчеркиваем важность этих полей в контексте интеграции и настоятельно рекомендуем использовать их в каждой записи.

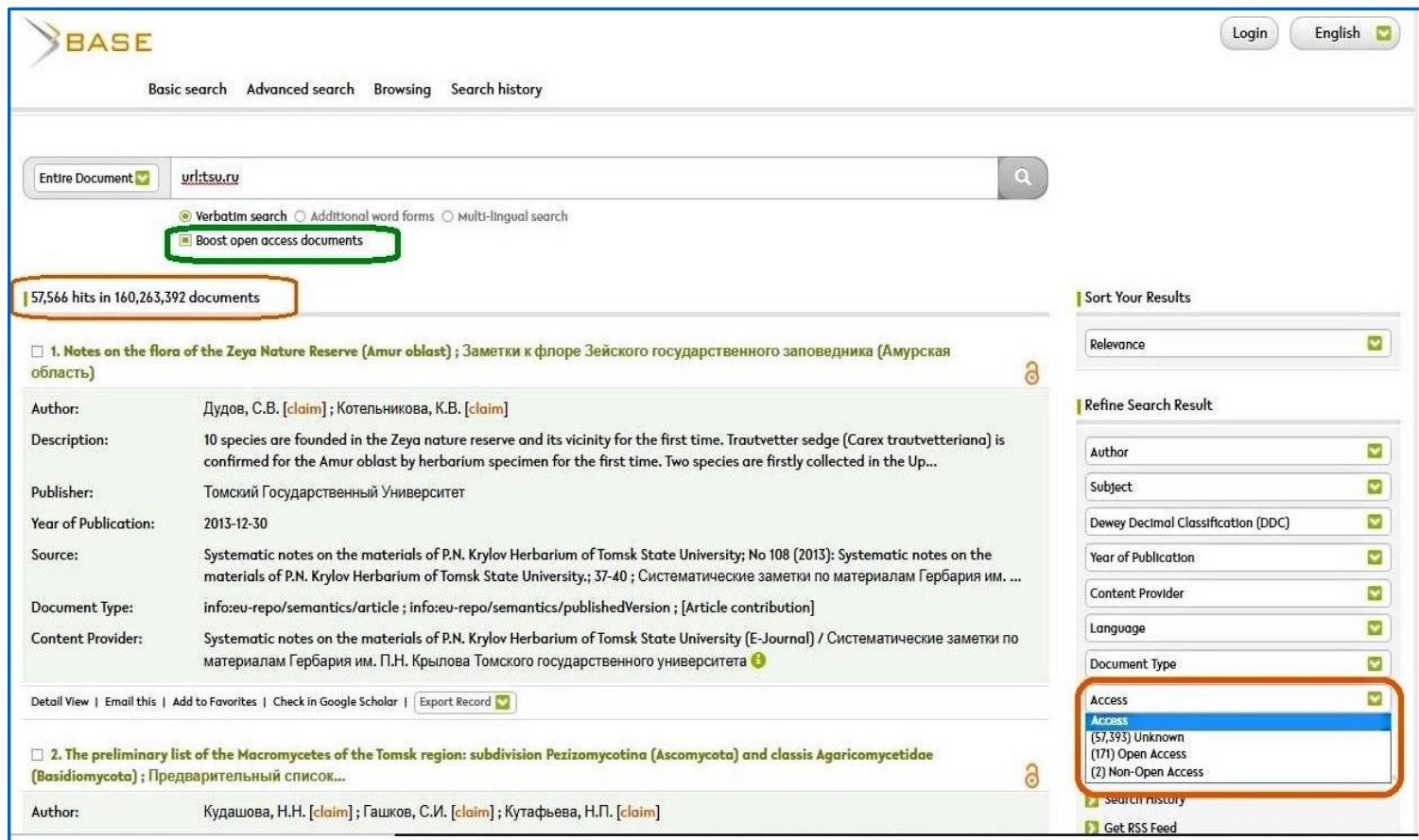
Рекомендация: здесь, в отдельном повторе поля, должна содержаться информация о правах доступа к ресурсу. Эта информация крайне важна для пользователей репозиторий-агрегаторов, поскольку именно она сигнализирует о том, что можно немедленно получить полный контент (полный текст).

Обязательность поля обусловлена требованиями международных агрегаторов, которые группируют записи по значениям этого поля и используют эти значения для фильтрации результатов поиска. Отсутствие в записи этого поля с принятыми значениями ведет к тому, что всем записям присваивается **неопределенный статус** (если только репозиторий-поставщик не декларировал, что ВСЕ записи – в свободном доступе).

В другие повторы поля записывается другая информация, имеющая отношение к правам на объект репозитория, например,

<dc:rights>Creative Commons Attribution 4.0 International</dc:rights>

Иллюстрация: BASE – ресурсы свободного доступа из репозитория Томского университета (по формальному признаку)



BASE Login English

Basic search Advanced search Browsing Search history

Entire Document ☒ url:tsu.ru

☒ Verbatim search ☐ Additional word forms ☐ Multi-lingual search

☒ Boost open access documents

57,566 hits in 160,263,392 documents

☐ 1. Notes on the flora of the Zeya Nature Reserve (Amur oblast) ; Заметки к флоре Зейского государственного заповедника (Амурская область)

Author: Дудов, С.В. [claim] ; Котельникова, К.В. [claim]

Description: 10 species are founded in the Zeya nature reserve and its vicinity for the first time. Trautvetter sedge (Carex trautvetteriana) is confirmed for the Amur oblast by herbarium specimen for the first time. Two species are firstly collected in the Up...

Publisher: Томский Государственный Университет

Year of Publication: 2013-12-30

Source: Systematic notes on the materials of P.N. Krylov Herbarium of Tomsk State University; No 108 (2013): Systematic notes on the materials of P.N. Krylov Herbarium of Tomsk State University.; 37-40 ; Систематические заметки по материалам Гербария им. ...

Document Type: info:eu-repo/semantics/article ; info:eu-repo/semantics/publishedVersion ; [Article contribution]

Content Provider: Systematic notes on the materials of P.N. Krylov Herbarium of Tomsk State University (E-Journal) / Систематические заметки по материалам Гербария им. П.Н. Крылова Томского государственного университета

Detail View | Email this | Add to Favorites | Check in Google Scholar | Export Record

☐ 2. The preliminary list of the Macromycetes of the Tomsk region: subdivision Pezizomycotina (Ascomycota) and classis Agaricomycetidae (Basidiomycota) ; Предварительный список...

Author: Кудашова, Н.Н. [claim] ; Гашков, С.И. [claim] ; Кутафьева, Н.П. [claim]

Sort Your Results: Relevance

Refine Search Result

Author ☒

Subject ☒

Dewey Decimal Classification (DDC) ☒

Year of Publication ☒

Content Provider ☒

Language ☒

Document Type ☒

Access ☒

Access

(57,393) Unknown

(171) Open Access

(2) Non-Open Access

Search history

Get RSS Feed

Рекомендации по заполнению dc:rights в части информации о правах доступа

Значения должны браться из перечня фиксированных значений во избежание неверной интерпретации.

Мы рекомендуем использовать значения из словаря COAR Controlled Vocabulary for Access Rights (http://vocabularies.coar-repositories.org/documentation/access_rights/).

Словари COAR – многоязычные, их используют OpenAIRE (с 4-й версии) и CORE; BASE распознает термины этих словарей в большинстве случаев. Словари разрабатываются Confederation of Open Access Repositories – организацией, которая занимается проблемами интероперабельности репозиториев. Большое внимание уделяется этой проблеме в недавно принятом документе “COAR Strategy 2019-2021 and Work Plan 2020”.

Контроль значений информации о правах доступа к ресурсам лежит в русле работ по стандартизации обмена данными между разными системами и сервисами, например, в Open Discovery Initiative. Включение формализованной информации о доступе к ресурсу – самый надежный способ довести эти сведения до конечного пользователя.

Допустимые значения



open access - означает свободный доступ любых пользователей к метаданным и полным текстам)

restricted access - означает наличие ограничений в доступе, например, доступ только для сотрудников организации

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

metadata only access – означает доступ к метаданным, но не к полному тексту.

Русскоязычные эквиваленты:

открытый доступ

доступ с эмбарго

ограниченный доступ

доступ только к метаданным

Если в репозитории – участнике НОРА вводятся русскоязычные термины, то при генерации контента НОРА в oai_dc они будут заменены соответствующими англоязычными.

Рекомендация: вводить в это поле одно из значений словаря COAR Resource Type (http://vocabularies.coar-repositories.org/documentation/resource_types/)

В международных агрегаторах это поле является обязательным. Тип ресурса определяет его отнесение к крупной категории, в пределах которой ведется поиск или которая служит фильтром результатов поиска. Особенно важно это для OpenAIRE, где поиск по умолчанию ведется в категории “Publications”, куда попадают ресурсы категории “Text”, наиболее развитой категории первого уровня (book, journal, journal article, thesis и другие). Если поле отсутствует, все записи попадают в категорию «Другие».

Иерархический перечень терминов словаря COAR приведен в Приложении 1 к «Методическим рекомендациям для разработчиков ОР по организации работы репозиторий: корректное создание метаданных».



National aggregator of open access repositories of Russian Universities

Национальный агрегатор открытых репозиторий российских
университетов (НОРА)

INSTITUTIONAL REPOSITORY AGGREGATOR

OPENAIRE 3.0 (OA, FUNDING)

OAI-PMH: <https://rour.neicon.ru/oai/request> →

Countries: Russian Federation

Publications (179,956)	+
Research Data (12)	+
Software (0)	+
Other Research Products (315,799)	+
Organizations (1)	+
Related Content Providers	+
Statistics	+
Metrics	+

Индивидуальный или коллективный автор, создатель.

Следует учитывать, что если в репозитории используются детализированные поля contributor (например, dc.contributor.editor), то для сборщика метаданных по OAI-PMH в формате oai_dc они будут представлены с метками базового dc – dc:contributor; Но: ***dc.contributor.author будет представлено как dc:creator***

Многие авторы работ, размещаемых в репозиториях, имеют идентификатор ORCID, широко используемый в международной практике. Для ввода ORCID автора мы рекомендуем один из двух вариантов на выбор репозитория-участника:

1) записать после имени автора в это поле в виде ссылки, разрешающейся в ORCID автора, заключенной в квадратные скобки

Пример: <dc:creator>Robinson, Neil[<https://orcid.org/0000-0002-0893-2190>]</dc:creator>

2) если такой вариант неудобен для держателя репозитория, то это поле может содержать только имена авторов; тогда ORCID следует включить в дополнительное поле dc.creatorORCID

dc:creatorORCID



Дополнительное поле *(не является обязательным)*.

Используется для записи ORCID автора при его имени и фамилии, если ORCID не был записан в поле автора.

Пример: <dc:creator>Robinson, Neil[<https://orcid.org/0000-0002-0893-2190>]</dc:creator>

Запись, содержащая ORCID, позволит репозиториям-агрегаторам выделить его по формальному признаку и занести в отдельное поисковое поле.

Обязательность распространяется на идентификатор, однозначно идентифицирующий объект в репозитории, URI. Именно он используется как ссылка, ведущая на ресурс в репозитории.

Поле может использоваться также для представления информации о других идентификаторах ресурса, например, обозначение записи в репозитории организации или идентификатор во внешних системах (Scopus, Web of Science, elibrary.ru, ...). Но для внешних систем важен именно идентификатор URI.

Особое место отводится цифровому идентификатору объекта – DOI, который сейчас становится практически универсальным средством поиска научных публикаций и основой ссылочного аппарата. В базовом DC отдельного поля для DOI нет, в квалифицированном его обычно записывают как dc.identifier.doi. Такова сложившаяся практика.

При этом предпочтительный способ записи DOI - в виде адреса, разрешающегося в ресурс, имеющий этот DOI:

<https://doi.org/10.1103/PhysRevB.37.785>

Это позволит внешним системам распознать DOI и занести значение в поисковое поле.

Решение команды НОРА для представления DOI: дополнительное поле dc:DOI

Доводы в пользу такого решения:

- 1) Важное поисковое поле будет представлено отдельным элементом, его не придется извлекать специальными программными средствами из поля идентификатора.
- 2) DOI не будет отождествляться с идентификатором записи в репозитории; в ряде случаев DOI относится к содержательно другому объекту, например, если в репозитории размещена авторская рукопись статьи. Некоторые репозитории получают DOI DataCite на записи в репозитории, а DOI опубликованной статьи приводят как DOI связанного объекта.

Пример – для записи из репозитория Apollo Кэмбриджского университета.

dc.title	A deep learning approach to bilingual lexicon induction in the biomedical domain
dc.type	Journal Article
dc.date.updated	2018-07-10T08:04:22Z
dc.language.rfc3066	en
dc.rights.holder	The Author(s)
dc.identifier.doi	10.17863/CAM.25260
rioxterms.versionofrecord	10.1186/s12859-018-2245-8

Международные агрегаторы рекомендуют обязательное использование этого поля: оно важно для поиска и вывода его результатов.

Обычно заполнение этого поля не вызывает вопросов. Следует только учитывать, что если в репозитории используются детализированные поля (например, `dc.title.alternative`), то для сборщика метаданных по OAI-PMH в формате `oai_dc` они будут представлены с метками базового `dc` – `dc:title`.

Подчеркнем, что в это поле вводится заглавие **только основного объекта описания**. В некоторых репозиториях сюда вводят названия объекта, включающего основной: название журнала для статьи.

dc:date

Обязательность этого поля не вызывает вопросов и соблюдается по факту: оно содержится во всех записях репозитория.

Для заполнения рекомендуется использовать стандарт ISO 8601 Date and Time Format

Развитие структур метаданных в международных агрегаторах - BASE

BASE пока сохраняет в качестве формата сбора метаданных базовый набор DC.

Правила подготовки метаданных для передачи в BASE кратко описаны в документе Golden Rules for Repository Managers (https://www.base-search.net/about/en/faq_oai.php).

В нем изложены некоторые приемы формализованного представления информации, для которой нет отдельных полей в структуре DC. В частности, предлагается включение ORCID авторов в поле dc:creator. Отмечается, что при обработке метаданных в них программно отыскиваются формализованно оформленные DOI и ORCID независимо от того, в каких элементах метаданных они присутствуют.

Развитие структур метаданных в международных агрегаторах - OpenAIRE

Рекомендации по подготовке метаданных для предоставления их по протоколу OAI-PMH изложены в нескольких версиях руководств (Guidelines). Действующая версия, принятая в 2013-м году, – третья. Ей на смену идет опубликованная в ноябре 2018 года четвертая версия, внедрение ее уже начато.

В третьей версии применяется стандартный формат Dublin Core; для структурированного представления некоторых важных данных (например, данных о финансирующих организациях) предлагается использовать значения стандартных полей DC, оформленных по специальному синтаксису, который позволяет разобрать эти значения на элементы для хранения и поиска.

В четвертой версии сделан важный шаг в направлении усложнения структуры метаданных с целью адекватного представления различных данных, прежде вводимых недифференцированно в поля DC широкой семантики. Для этого применяется три пространства имен (namespaces): DC, DataCite и OpenAIRE. Так, отдельные поля предусмотрены для ввода данных о проектах (oaire:fundingReference), альтернативных идентификаторах (datacite:alternateIdentifier) и других.

Развитие структур метаданных в международных агрегаторах - CORE

Структура метаданных репозитория описана в документе по адресу <https://core.ac.uk/ref-audit/#recommendations>. Эти рекомендации в некоторой степени ориентированы на репозитории университетов Великобритании, поэтому для представления метаданных используется формат RIOXX (<http://www.rioxx.net/>), в котором выделены специальные поля для записи ORCID авторов, DOI опубликованной версии статьи, размещенной в репозитории, и детализированного представления дат (дата принятия рукописи в печать, дата публикации). В этом формате используются как метки DC, так и метки rioxxterms для специальных полей.

Формат совместим с форматом OpenAIRE версии 3: разработана таблица соответствия полей двух форматов – Crosswalk from RIOXX 2/0 to OpenAIRE 3.0. (https://www.rioxx.net/mappings/crosswalk_rioxx_2_0_openaire_3_0/)

Развитие структур метаданных - тенденции

Развитие структур метаданных репозиторий идет в направлении бОльшего разнообразия и дифференциации элементов. Это связано с повышением роли репозиторий в структуре открытой науки, с размещением научных публикаций, подготовленных при поддержке фондов, в репозиториях как альтернативе публикациям в журналах открытого доступа.

Учитывая эти тенденции, мы можем выделять в своих внутренних структурах метаданных отдельные элементы, чтобы впоследствии передавать их в детализированных обменных форматах.

СПАСИБО ЗА ВНИМАНИЕ

Наталия Николаевна Литвинова
nlit@neicon.ru