

# Автоматическое пополнение репозитория организации

Денис Викторович Косяков

Scopus®



ORCID

Connecting Research  
and Researchers



Crossref

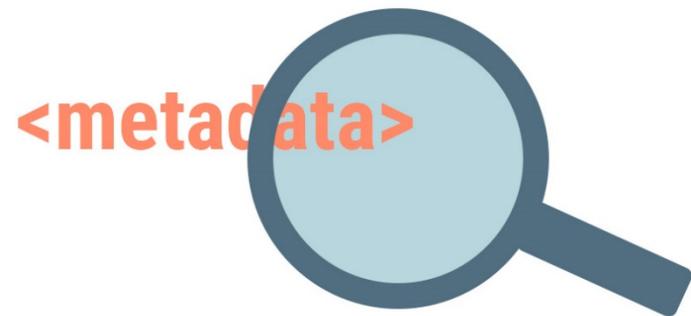


SHERPA/ROMEO



НАУЧНАЯ ЭЛЕКТРОННАЯ  
БИБЛИОТЕКА

LIBRARY.RU



Где взять метаданные публикаций  
моих авторов?

## Источники данных

- API доступен в рамках национальной подписки
- ресурсы для разработчиков <https://dev.elsevier.com/>
- Scopus Search API – поиск по идентификатору аффилиации

*http://api.elsevier.com/content/search/scopus?view=COMPLETE&  
query=AF-ID(60107593) AND PUBYEAR = 2020*

- Abstract Retrieval API – извлечение полной записи по идентификатору

*http://api.elsevier.com/content/abstract/eid/  
2-s2.0-85081419786?view=FULL*

- по умолчанию представления COMPLETE и FULL недоступны, но их могут разрешить по письму в техподдержку. В стандартных представлениях информация (об авторах) неполна!
- NB! – по условиям использования API «**abstract** may **not** be displayed publicly»

# Источники данных

## Public API

- позволяет извлекать публичные данные профилей
- список публикаций:

*[https://pub.orcid.org/v2.0/  
0000-0002-0495-9898/works](https://pub.orcid.org/v2.0/0000-0002-0495-9898/works)*

- формат записи
- источники пополнения



<https://members.orcid.org/api/about-public-api>

## Экспорт XML (Science Index для организаций)

## API (лицензируется отдельно)

The screenshot shows the 'СПИСОК ПУБЛИКАЦИЙ ОРГАНИЗАЦИИ' page on the Science Index website. The page is for the 'ГОСУДАРСТВЕННАЯ ПУБЛИЧНАЯ НАУЧНО-ТЕХНИЧЕСКАЯ БИБЛИОТЕКА СО РАН' in Novosibirsk. It features a search interface with various filters and a list of publications.

**ПАРАМЕТРЫ**

Подразделение: [dropdown]

▼ ТЕМАТИКА  
▼ ЖУРНАЛЫ  
▼ ОРГАНИЗАЦИИ  
▼ АВТОРЫ  
▼ ГОДЫ  
▼ ТИП ПУБЛИКАЦИИ  
▼ УЧАСТИЕ В ПУБЛИКАЦИИ

Выбрать: [dropdown]

Показывать: [dropdown]

- учитывать публикации, извлеченные из списков цитируемой литературы  
 - объединять оригинальные и переводные версии статей и переиздания книг

Сортировка: [dropdown]    Порядок: [dropdown]    [Очистить]    [Поиск]

**ИНСТРУМЕНТЫ**

- Следующая страница
- Выделить все публикации на этой странице
- Снять выделение
- Добавить выделенные публикации в подборку: [dropdown]
- Добавить все публикации организации в указанную выше подборку
- Вывести список публикаций, ссылающихся на публикации организации
- Анализ публикационной активности организации
- Публикационная активность организации в инфографике
- Вывести на печать список публикаций организации
- Экспорт списка публикаций организации в формате XML
- Создать описание публикации на основе информации из ссылки
- Структурные подразделения организации
- Список сотрудников организации

**Всего найдено 2597 публикаций с общим количеством цитирований: 5314. Показано на данной странице: с 1 по 100.**

№	Публикация	Цит.
1.	<b>РУКОПИСНЫЕ И СТАРОПЕЧАТНЫЕ КНИГИ НАУЧНОЙ БИБЛИОТЕКИ ХАБАРОВСКОЙ ДУХОВНОЙ СЕМИНАРИИ</b> Бороздин А.Ю., Шилова И.А., Кудин А.А. Библиография. Научный журнал по библиографоведению, книговедению и библиотековедению. 2020. № 1 (426). С. 78-87.	0
2.	<b>АНАЛИТИЧЕСКИЕ СИСТЕМЫ В ОЦЕНКЕ БИБЛИОТЕЧНЫХ САЙТОВ</b> Ударица О.М. Библиоковедение. 2020. Т. 69. № 1. С. 31-42.	0
3.	<b>БИБЛИОТЕКА "ОБЩЕСТВА ВЗАИМНОГО ВСПОМОЖЕНИЯ ПРИКАЗЧИКОВ В Г. ТОМСКЕ": ИСТОРИЯ И ОСНОВНЫЕ ПОКАЗАТЕЛИ ДЕЯТЕЛЬНОСТИ (1894-1913)</b> Тимофеева Ю.В. Библиоковедение. 2020. Т. 69. № 1. С. 85-99.	0
4.	<b>НАВИГАТОР ПО ВТОРИЧНЫМ ИСТОЧНИКАМ РЕГИОНАЛЬНОЙ ТЕМАТИКИ КАК КОМПОНЕНТ ИНФОРМАЦИОННОГО ОБЕСПЕЧЕНИЯ НАУЧНЫХ ИССЛЕДОВАНИЙ (НА ПРИМЕРЕ ТЕМЫ "ТЕАТР, МУЗЫКА, ТАНЦЫ")</b> Бусыгина Т.В., Мандригина Л.А., Рыкова В.В. Информационные ресурсы России. 2020. № 1 (173). С. 13-18.	0

- GetAuthorIdExt – список id авторов организации
- GetAuthor2 – данные автора
- GetItemId – список идентификаторов публикаций
- GetItem2 – библиографическая запись

в ответе XML

Источники  
данных



Данные о публикациях получили, а как насчет полных текстов?

## Источники данных

Кроме метаданных самой публикации многие издатели размещают ссылку на полный текст

- headers *Accept* = "*application/citeproc+json*"
- *http://dx.doi.org/10.20913/1815-3186-2020-1-12-2*
- link  
*URL : https://www.bibliosphere.ru/jour/article/viewFile/1599/1623*  
*content-type : application/pdf*  
*content-version : vor*  
*intended-application : text-mining*



Имеем ли мы право на публикацию  
полного текста в открытом доступе  
и, если «да», то в какой версии?

# Проверка ПОЛИТИКИ ИЗДАТЕЛЯ

Sherpa Romeo PUBLIC TEST RELEASE

About Search Statistics Support Us Contact Admin

## Journal of Medical Internet Research

This version of Sherpa Romeo is in its public testing period. Use our survey or email [help@jisc.ac.uk](mailto:help@jisc.ac.uk) to give feedback or report a bug.

### Publication Information

Title	Journal of Medical Internet Research (JMIR) <small>(English)</small>
ISSNs	Electronic: 1438-8871
URL	<a href="http://www.jmir.org/">http://www.jmir.org/</a>
Publishers	JMIR Publications [Commercial Publisher]
Listed in DOAJ	Yes <small>[<a href="https://doaj.org/toc/1438-8871">https://doaj.org/toc/1438-8871</a>]</small>

### Publisher Policy

Submitted Version Accepted Version Published Version

The following Open Access route is permitted for this version:

Location	Any Website This Journal
Embargo	No Embargo
Licence	CC BY 4.0
Conditions	Published source must be acknowledged with citation Must link to published article

For more information, please see the following links:

- [Copyright Information](#)



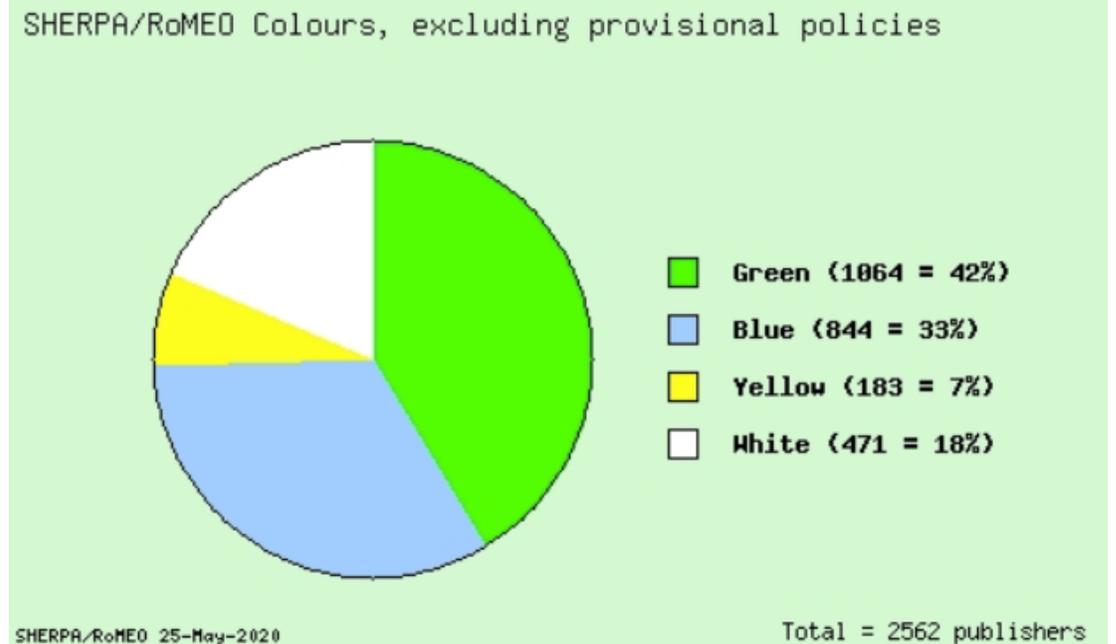
• [http://www.sherpa.ac.uk/romeo/api29.php?ak=\\*\\*\\*\\*\\*&issn=14388871](http://www.sherpa.ac.uk/romeo/api29.php?ak=*****&issn=14388871)

```
id : 960
name : Journal of
      Medical Internet
      Research
homeurl : http://
         www.jmir.org/
preprints : preprints
postprints : postprints
pdfversion : pdfversion
conditions : conditions
paidaccess : paidaccess
romeocolour : green
dateadded : 2011-02-18 11:22:43
dateupdated : 2016-01-26 11:56:23
```

# Статистика SHERPA / RoMEO

RoMEO colour	Archiving policy	Publishers	%
<a href="#">green</a>	Can archive pre-print and post-print	1064	42
<a href="#">blue</a>	Can archive post-print (ie final draft post-refereeing)	844	33
<a href="#">yellow</a>	Can archive pre-print (ie pre-refereeing)	183	7
<a href="#">white</a>	Archiving not formally supported	471	18

Summary: **82%** of publishers on this list formally **allow** some form of self-archiving.





**DSPACE**

Как разместить публикации  
в репозитории?

# DSpace 7 API

## API полностью переработан в версии 7.

- позволяет отправлять новые документы (POST /api/submission/workspaceitems)
- вносить исправления (PATCH api/submission/workspaceitems/[id])

в конце надо

- отдать указатель на новый документ рабочему процессу (POST api/workflow/workflowitems -H Content-Type: "text/uri-list" -data '/api/submission/workspaceitems/[id]')
- найти задачу (POST api/submission/pooltasks/search/findByUser?uuid=335647b6-8a52-4ecb-a8c1-7ebabb199bda)
- взять задачу (POST api/workflow/pooltasks/[id] -H Content-Type: "x-www-form-urlencoded")
- извлечь данные задачи (POST api/submission/claimedtasks/search/findByUser=335647b6-8a52-4ecb-a8c1-7ebabb199bda)
- завершить задачу (POST api/workflow/claimedtasks/[id]?submit\_approve=true)

Итак, весь процесс

- найти и извлечь метаданные

- найти и извлечь полный текст (при наличии)

- определить права на публикацию полного текста

- отправить данные и полный текст в репозиторий

# Спасибо за внимание!

Денис Викторович Косяков

ГПНТБ СО РАН

[kosyakov@spsl.nsc.ru](mailto:kosyakov@spsl.nsc.ru)