

OpenAlex

katalog odprtih raziskovalnih objav

Vladimir Batagelj

UP IAM in FAMNIT, IMFM, UL FMF

OZUP

po Zoomu, 17. april 2025

1 OpenAlex



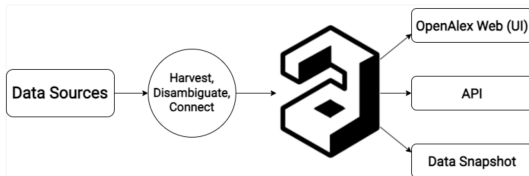
Vladimir Batagelj: vladimir.batagelj@fmf.uni-lj.si

Current version of slides (April 17, 2025 at 00:32): [slides](#) [PDF](#)

<https://github.com/bavla/OpenAlex>

OpenAlex

OpenAlex je popolnoma odprt katalog globalnega raziskovalnega sistema [1]. Ime je dobil po starodavni Aleksandrijski knjižnici. Izdelala ga je neprofitna organizacija **OurResearch**, ki jo podpira Arcadia. OpenAlex je bil predstavljen januarja 2022.



Ponuja dostop preko uporabniškega vmesnika ali brezplačnega API-ja. Na svoj računalnik lahko prenesemo tudi posnetek celotnega podatkovja. Je nadomestek storitve Microsoft Academic Graph (MAG), ki je bila umaknjena 31. decembra 2021 [2].

Francosko ministrstvo za visoko šolstvo in raziskave **sodeluje** z OpenAlexom pri razvoju popolnoma odprtega bibliografskega orodja. CNRS je **prekinil naročenost** na bibliografske storitve Scopus.

Wikipedia/OpenAlex

Raziskovalci, financerji in organizacije po vsem svetu se pri iskanju, izvedbi in upravljanju svojih raziskav zanašajo na bibliografske baze. Desetletja so te informacije zagotavljale le tržne storitve (Web of Science, Scopus, itd. – knjižnice so stale okrog milijarde dolarjev letno). Te storitve izključujejo dela z nekaterih tematskih in zemljepisnih področij. Tudi naročniki imajo omejeno uporabo podatkov.

OpenAlex zajema več kot dvakrat več znanstvenih del kot vodilne tržne storitve.

Vrste bibliografskih enot

V. Batagelj

OpenAlex

References



OpenAlex temelji na 7 vrstah ent: dela **W**(ork), avtorji **A**(uthor), viri **S**(ource), ustanove **I**(nstitution), pojmi **C**(oncept), izdajatelji **P**(ublisher) in podporniki **F**(under).

OpenAlex rešuje več za analizo bibliografskih podatkov pomembnih vprašanj:

- 1 **identifikacija** bibliografskih enot (ID-ji) [3]
- 2 brezplačen dostop do podatkov in možnost deljenja izpeljanih podatkov (**Prenos na uporabnikov računalnik**),
- 3 izboljševanje podatkovja s sodelovanjem uporabnikov (**Oddaj zahtevo**).

- Storitve OpenAlex <https://openalex.org/>
- Avtor z ID <https://openalex.org/A5001676164>
- Delo z DOI <https://api.openalex.org/works/https://doi.org/10.1007/s11192-012-0940-1>
- Delo z ID <https://openalex.org/W2083084326>
- Ime ustanove
<https://api.openalex.org/institutions?search=imfm>
- Ustanova z ID <https://openalex.org/institutions/I4210106342>

OpenAlex

Osebnna stran

OpenAlex

V. Batagelj

OpenAlex

References

```

<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<html>
  <head><title>Osebnna stran: Vladimir Batagelj</title></head>
  <body>
    <table><tr>
      <td> &nbsp; &nbsp; </td>
      <td> <h1>Vladimir Batagelj</h1> Professor emeritus,
      <a href="http://www.uni-lj.si/">University of Ljubljana</a><br/>
      <a href="http://www.fmf.uni-lj.si/si/">FMF - Department of
      Mathematics</a><br/> &nbsp; &nbsp; and<br/>
      <a href="http://www.imfm.si/">IMFM - Institute of Mathematics,
      Physics and Mechanics</a><br/>
      Department for theoretical computer science<br/>
      Jadranska 19, 1000 Ljubljana, Slovenia<br/>
      e-mail: vladimir.batagelj@fmf.uni-lj.si<br/> &nbsp; &nbsp; and<br/>
      University of Primorska, <a href="http://www.iam.upr.si/en/">
      Andrej Marušič Institute</a>, Koper
    </td></tr></table><br/>
    <iframe src="https://openalex.org/authors/A5001676164" width="100%"
      height="500" style="border:1px solid black;"></iframe>
  </body>
</html>

```

Prikaži osebno stran

OpenAlex

Search, filter, select

OpenAlex

V. Batagelj

OpenAlex

References

The OpenAlex API is available at <https://api.openalex.org>. Its response is returned in JSON format.

Using **search** we can search for a given search text across titles, abstracts, and full-text. Using a **filter** we can limit our search to units satisfying given conditions. Using **select** we can select data fields that will appear in results.

List of work IDs with titles

The OpenAlex API uses paging – the list data are provided by pages. The **basic paging** (up to 10 000 units) is based on two parameters `page` and `per_page`). **Cursor paging** is a bit more complicated than basic paging, but it allows us to access as many records as we like.

We developed an R package `OpenAlex2Pajek` for constructing a collection of Pajek bibliographic networks on selected topic from **OpenAlex**.

OpenAlex

Using API from program

OpenAlex

V. Batagelj

OpenAlex

References

Here is an R code using the OpenAlex API

```

setwd(wdir <- "C:/work/OpenAlex/API")
library(httr); library(jsonlite)
wd <- GET("https://api.openalex.org/works",
  query = list(
    search="handball",
    filter="publication_year:2015",
    select="id,title",
    page="2", per_page="200"))
names(wd)
wc <- fromJSON(rawToChar(wd$content))
names(wc); names(wc$meta)
wc$meta$count; str(wc$results)
    
```

It returns the second page (with up to 200 entries) on works on handball published in the year 2015. Only information about works ID and title is returned.

As a data analyst, I miss short names for individual units (Garfield, journal abbreviations, etc.).

Person names are not structured (First, Mid, Last).

The problem of author countries – my example of an "extinct" country. [W2033820728](#), [W2059649701](#) ([JSONview](#)) Click API, see institutions

Missing relations to derived works (preprint – published, translation, book edition, etc.).

To ensure the OpenAlex longevity – UNESCO?

References I

OpenAlex

V. Batagelj

OpenAlex

References

1. Priem, J., Piwowar, H. & Orr, R. OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. *arXiv preprint arXiv:2205.01833* (2022).
2. Chawla, D. S. Massive open index of scholarly papers launches. *Nature* (2022).
3. DiRROS Data. *Trajni identifikatorji digitalnih in ne-digitalnih objektov*. <https://dirrosdata.ctlk.uni-lj.si/repozitoriji/trajni-identifikatorji/>. 2022.