



OpenAlex

An open & complete index of the
global research system

- Scientific Knowledge Graphs (SKGs) are vital, but closed and broken
- OpenAlex is an open SKG with excellent coverage
- You should use OpenAlex

**Scientific Knowledge Graphs (SKGs)
are essential infrastructure.**

What are SKGs?

- Web of Science
- Scopus
- Google Scholar?
- Microsoft Academic Graph
- Dimensions
- Crossref, COCI
- OpenAIRE

Pay-to-view SKGs are badly broken.

Pay-to-view SKGs are broken

Can't afford them

Can't share them

Can't build on them.

Pay-to-view SKGs are broken

Can't afford them

Can't share them

Can't build on them.

We can't afford toll-access SKGs

Who can't afford them?

We can't afford toll-access SKGs

We're **systematically excluding** less wealthy regions from understanding and managing research.

We can't afford toll-access SKGs

I wonder if that's why
Scopus/WoS have <8%
non-english articles? 🤔

Pay-to-view SKGs are broken

Can't afford them

Can't share them

Can't build on them.

We can't share toll-access SKGs

- transparency
- reproducibility
- integrity

Pay-to-view SKGs are broken

Can't afford them

Can't share them

Can't build on them.

OpenAlex is an open SKG.

OpenAlex is an open SKG

Open source **code**: MIT

Open **data**: CC0

Usably open: API, web UI, snapshot

Sustainably open: PCSI nonprofit

OpenAlex API: filter

https://api.openalex.org/works?filter=institutions.ror:https://ror.org/01cwqze88

+ view source

```
{
- meta: {
  count: 320955,
  db_response_time_ms: 70,
  page: 1,
  per_page: 25
},
- results: [
  - {
    id: "https://openalex.org/W3175558339",
    doi: "https://doi.org/10.1080/21645515.2021.1908030",
    title: "Prediction of serum HIV-1 neutralization titers of VRC01 in HIV-uninfected Antibody Mediated Prevention",
    display_name: "Prediction of serum HIV-1 neutralization titers of VRC01 in HIV-uninfected Antibody Mediated Pre",
    publication_year: 2022,
    publication_date: "2022-12-31",
    - ids: {
      openalex: "https://openalex.org/W3175558339",
      doi: "https://doi.org/10.1080/21645515.2021.1908030",
      mag: "3175558339",
      pmid: "https://pubmed.ncbi.nlm.nih.gov/34213402"
    },
    - host_venue: {
      id: "https://openalex.org/V2483136528",
      issn_l: "2164-5515",
      - issn: [
        "2164-5515",
        "2164-554X"
      ],
      display_name: "Human Vaccines & Immunotherapeutics",
```


OpenAlex snapshot

The screenshot shows a web browser window with the address bar displaying `openalex.s3.amazonaws.com/browse.html`. The browser's toolbar includes navigation icons, a star for bookmarks, and various extension icons. The AWS S3 Explorer interface is visible, with a blue header bar containing the text "AWS S3 Explorer" and a button labeled "openalex". To the right of the header are controls for "Hide folders?", "Folder", "Bucket", a settings gear icon, and a notification badge with the number "5". Below the header, there is a "Show 50 entries" dropdown and a "Search:" input field. The main content area displays a table of objects in the bucket.

Object	Last Modified	Timestamp	Size
data/			
browse.html	22 days ago	2022-05-18 13:09:34	32 KB
LICENSE.txt	5 months ago	2022-01-02 14:26:34	7 KB
README.txt	5 months ago	2022-01-03 13:32:18	301 Bytes
RELEASE_NOTES.txt	a month ago	2022-05-12 16:26:59	2 KB

OpenAlex web interface



Works ▾ frogs X ▾

Home About

20 results   Sort by Relevance ▾  

[Effects of extracts of seed and leaf of Piper guineense on skeletal muscle activity in rat and frog.](#)
F. V. Udoh, T. Y. Lot, et al. - *Phytotherapy Research*, 1999
Cited by 27 [Fulltext via publisher](#)

[Frog Consumption Pattern in Ibadan, Nigeria](#)
F A. Akinyemi, Efenakpo Dean Ogaga - *Journal for Studies in Management and Planning*, 2015
Cited by 2

[Mycoflora, Mycotoxin Contamination and Proximate Mineral Composition of Smoke-Dried Frog \(Aubria sp.\) \(Konko\) Sold in Ibadan, Oyo State, Nigeria](#)
Bukola Christianah Adebayo-Tayo, Folahanmi Adeyemi, et al. - *Turkish Journal of Agriculture: Food Science...*, 2015
Cited by 4 [Fulltext via publisher](#)

[Cryptosporidium in Wild Frogs \(Rana spp \) Consumed by humans in Kaduna State Nigeria](#)
G.S.N. Kia, Blessing Iveren Ukuma, et al. - *Online Journal of Public Health Informatics*, 2017
Cited by 1 [Fulltext via publisher](#)

[Anurans as Intermediate and Paratenic Hosts of Helminth Infections in the Rainforest and Derived Savanna Biotopes of Southern Nigeria](#)
A.A. Imasuen, Habibat J. Ozemoka, et al. - *International Journal of Zoology*, 2012
Cited by 10 [Fulltext via publisher](#)

Work: journal article
Effects of extracts of seed and leaf of Piper guineense on skeletal muscle activity in rat and frog.
[Phytotherapy Research](#) (1999)



3 Authors ([show details](#))
[F. V. Udoh](#), [T. Y. Lot](#), [V. B. Braide](#)

5 Concepts ([show all](#))
[Biology](#), [Chemistry](#), [Traditional medicine](#)

[27 incoming citations view](#)
[6 outgoing references](#)
[20 related works](#)

[FULLTEXT VIA PUBLISHER](#)



- Concepts** ▾
- Venues** ▾
- Institution** ▾
- Author** ▾
- Fulltext Available** 
 - False 33
 - True 20
- Publisher** ▾
- Country** 
 - Unknown 25,071
 - USA 5,076
 - China 1,234
 - Japan 1,233
 - UK 1,140
 - Nigeria 20

MORE
- Type** ▾

**OpenAlex can be open because
it's super cheap to make.**

Why is OpenAlex so cheap?

Open MAG data

Open Source ML

Open Access literature

Open PID graphs

Open PID graphs

Works: Crossref: DOI, citations

Authors: ORCID

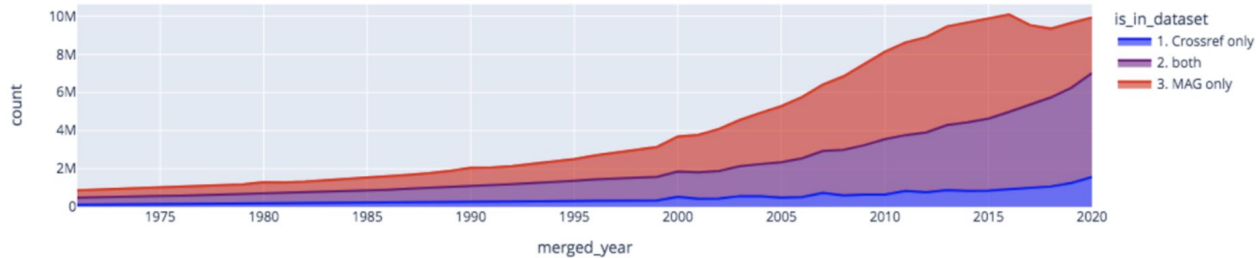
Journals: ISSN

Institutions: ROR

Concepts: Wikidata

DOI covers more and more.

Works per year, MAG and Crossref



MAG-only works per year



**OpenAlex has broader coverage
than any other SKG.**

SKG coverage (and Google Scholar)

	Works	Citations	Authors	Venues	Institutions
Web of Science (core)	82M	1.8B	17M	24k	11k
Scopus	82M	1.7B	17M	40k	80k
Google Scholar	390M	?	4M?	?	-
MAG	204M	1.7B	206B	49k	27k
Dimensions	148M	1.6B	28k	74k	100k
Crossref/COCI	135M	1.3B	4.5M	104k	1k
OpenAlex	236M	1.8B	243M	124k	108k

SKG coverage sources

- * <https://link.springer.com/article/10.1007/s11192-018-2958-5>
- * <https://link.springer.com/article/10.1007/s11192-021-03948-5>
- * <https://www.crossref.org/06members/53status.html>
- * <https://api.openalex.org>
- * <https://api.crossref.org/works>
- * <https://app.dimensions.ai/discover/publication>
- * <https://clarivate.libguides.com/librarianresources/coverage>
- * https://api.crossref.org/works?facet=ror-id:*
- * queries on closed-source WoS database
- * <https://www.mdpi.com/2304-6775/9/1/12>
- * <https://twitter.com/albertomartin/status/1534088434427604992>
- * <https://twitter.com/digitalsci/status/1534182383066525696>

OpenAlex coverage of the Global South

22k institutions

21M authors

22M works

1/3 of works are in
non-English language

[source: our API!](#)

```
...
188
189 def institutions_in_global_south():
190     total = 0
191     for country in GLOBAL_SOUTH_COUNTRIES:
192         country_code = countries.get(country).alpha2
193         r = requests.get(
194             f"https://api.openalex.org/institutions?filter=country_code:{country_code}&mailto:jason@ourresearch.org"
195         )
196         institution_count = r.json()["meta"]["count"]
197         print(
198             f"({institution_count}) institutions in {countries.get(country).name} ({country_code})"
199         )
200         total = total + institution_count
201     return total
202
203 def works_from_global_south():
204     total = 0
205     for country in GLOBAL_SOUTH_COUNTRIES:
206         country_code = countries.get(country).alpha2
207         r = requests.get(
208             f"https://api.openalex.org/works?filter=authorships.institutions.country_code:{country_code}&mailto:jason@ourresearch.org"
209         )
210         works_count = r.json()["meta"]["count"]
211         print(
212             f"({works_count}) works from {countries.get(country).name} ({country_code})"
213         )
214         total = total + works_count
215     return total
216
217 def authors_from_global_south():
218     total = 0
219     for country in GLOBAL_SOUTH_COUNTRIES:
220         country_code = countries.get(country).alpha2
221         r = requests.get(
222             f"https://api.openalex.org/authors?filter=last_known_institution.country_code:{country_code}&mailto:jason@ourresearch.org"
223         )
224         authors_count = r.json()["meta"]["count"]
225         print(
226             f"({authors_count}) authors from {countries.get(country).name} ({country_code})"
227         )
228         total = total + authors_count
229     return total
230
231
232
233 if __name__ == "__main__":
234     institutions_count = institutions_in_global_south()
235     print(f"({institutions_count},) institutions in the Global South")
236     # 22,073 institutions in the Global South
237
238     works_count = works_from_global_south()
239     print(f"({works_count},) works from the Global South")
240     # 22,978,988 works from the Global South
241
242     authors_count = authors_from_global_south()
243     print(f"({authors_count},) authors with last known affiliation in Global South")
```

OpenAlex connects five entity types:
works, **authors,** **venues,** **institutions,** & **concepts.**

OpenAlex entities

- **Works**: papers, datasets, software, etc.
- **Authors**: people who *create* works
- **Venues**: journals and repositories that *host* works
- **Institutions**: *affiliations claimed* by authors of works
- **Concepts**: what works are *about*

236M

DOI, PMID

OpenAlex Authors

243M

ORCID

OpenAlex Venues

108k

ROR

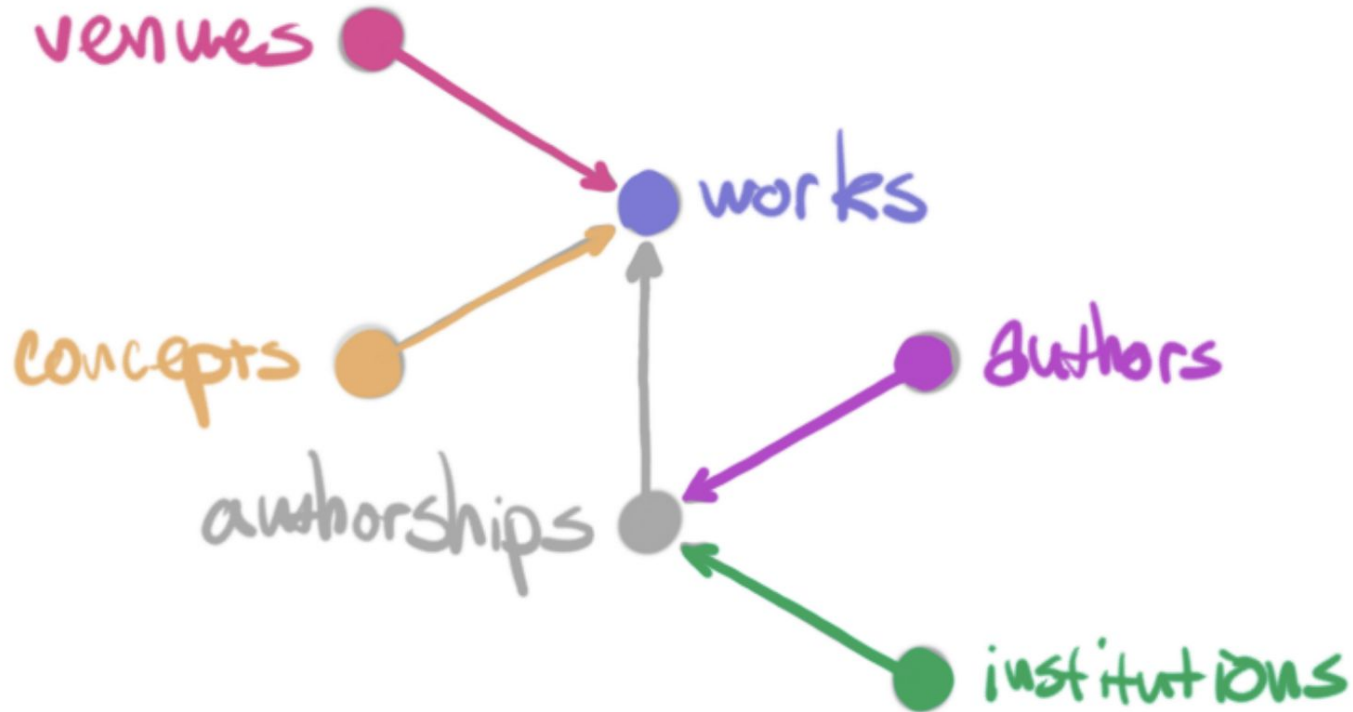
124k

ORCID

60k

Wikidata

Basic data model



OpenAlex has limitations.

OpenAlex limitations

- **Bias:** inherits from sources, esp Crossref and MAG
- **Research base:** not much, yet.
- **Coverage:** No funders or patents. Limited software, datasets, books.
- **Stability:** Data changing (improving) monthly

**OpenAlex is open and ready for use.
Go play with it!**