An open & complete index of the global research ecosystem
Today’s Agenda

Introducing OpenAlex
Why you should use it
  - major features
  - how it compares
How you can start using it
  - access methods
  - form & function of the database
Q&A
Scientific Knowledge Graphs (SKGs) are essential infrastructure.
- research discovery
- scientometrics
- research intelligence and assessment
What are SKGs?

- Web of Science
- Scopus
- Google Scholar?
- Microsoft Academic Graph
- Dimensions
- Crossref, COCI
- OpenAIRE
OpenAlex is an open SKG. And that’s a major feature.
Limitations of Pay-to-view SKGs

Their subscriptions are costly
Their (your) results cannot be shared
You can’t build on them
You inherit their exclusiveness
Their subscriptions are costly

- Pressure on budgets is intensifying at Universities (esp. library)
- Paywalls systematically exclude less wealthy regions
- After paying for subscription, your access is limited

OpenAlex is free, enabling equitable access across the globe
Their (your) results cannot be shared

- transparency in decision-making
- reproducibility of meta-research

Because OpenAlex is completely open, anyone can examine and replicate analyses and scenario-play factors influencing decisions.
You can’t build on them

Possibilities limited by closed databases & their licenses:
- access to full datasets
- commercial uses of data
- integrations with internal or external dashboards
- development of derivative tools

OpenAlex data and codes are under CCO, anyone can examine and use however they wish without lawyers
You inherit their exclusiveness

Exclusiveness criteria that create biases:
- must have an English abstract
- publication status
- theses/dissertations excluded
- type of peer review
- journal has relatively few citations
- geographical diversity of authors

OpenAlex does not apply indexing criteria so you can pick which data to include for your purposes
OpenAlex has broader coverage than any other SKG.
## SKG content coverage

<table>
<thead>
<tr>
<th>Data Source</th>
<th>Works</th>
<th>Citations</th>
<th>Authors</th>
<th>Venues</th>
<th>Institutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web of Science (core)</td>
<td>82M</td>
<td>1.8B</td>
<td>17M</td>
<td>24k</td>
<td>11k</td>
</tr>
<tr>
<td>Scopus</td>
<td>82M</td>
<td>1.7B</td>
<td>17M</td>
<td>40k</td>
<td>80k</td>
</tr>
<tr>
<td>Google Scholar</td>
<td>390M</td>
<td>?</td>
<td>4M?</td>
<td>?</td>
<td>-</td>
</tr>
<tr>
<td>MAG</td>
<td>204M</td>
<td>1.7B</td>
<td>206B</td>
<td>49k</td>
<td>27k</td>
</tr>
<tr>
<td>Dimensions</td>
<td>148M</td>
<td>1.6B</td>
<td>28k</td>
<td>74k</td>
<td>100k</td>
</tr>
<tr>
<td>Crossref/COCI</td>
<td>135M</td>
<td>1.3B</td>
<td>4.5M</td>
<td>104k</td>
<td>1k</td>
</tr>
<tr>
<td>OpenAlex</td>
<td>236M</td>
<td>1.8B</td>
<td>243M</td>
<td>124k</td>
<td>108k</td>
</tr>
</tbody>
</table>
OpenAlex indexes works & associated metadata

Sources [232k]
Concepts [65k]
Publishers [10k]
Institutions [102k]
Works [243M]
Funders [32k]
Authors [91M]

Sudden collapse of a mesopredator reveals its complementary role in mediating rocky reef regime shifts
Work (Article)

Published: 2018
Source: Proceedings of The Royal Society B: Biological Sciences
Authors: Jenn M. Burt (Simon Fraser University), M. Tim Tinker (University of California, Santa Cruz), Daniel K. Okamoto (Florida State University, Simon Fraser University), Kyle W. Demes (University of British Columbia, Simon Fraser University), Keith Holmes (Tula Foundation), Anne K. Salamon (Simon Fraser University)
Concepts: Trophic cascade, Mesopredator release hypothesis, Reef
Cited by: 76
Cites: 55
Related: 10

Abstract
While changes in the abundance of keystone predators can have cascading effects resulting in regime shifts, the role of mesopredators in these processes remains understudied. We conducted annual surveys of rocky reef communities that varied in the recovery of a keystone predator (sea otter, Enhydra lutris) and the mass mortality of a mesopredator (sunflower sea star, Pycnopodia helianthoides) due to an infectious wasting disease. By fitting a population model to empirical data, we show that sea otters had the greatest impact on the mortality of large sea urchins, but that Pycnopodia decline corresponded to a 311% increase in medium urchins and a 30% decline in kelp densities. Our results reveal that predator complementarity in size-selective prey consumption strengthens top-down control on urchins, affecting the resilience of alternative reef states by reinforcing the resilience of kelp forests and eroding the resilience of urchin barren states. We reveal previously underappreciated species interactions within a ‘classic’ trophic cascade and regime shift, highlighting the critical role of middle-level predators in mediating rocky reef state transitions.

Funders (4)
Locations (4)
Identifiers (4)
Some of the metadata available (docs.openalex.org/)

**Work**
- Bibliographical [title, publication date/year]
- Related PIDs [ID, DOI, PMID, PMCID, MAG,]
- Access [License, OA Color, OA version, OA Link]
- Authors & Institutions [#, Names, IDs, Corresponding]
- APC [List price, Paid, Paid: provenance]
- Repository locations [Name, Availability, Distinct #]
- Funder [Name, Grant ID]
- Citations [Referenced, Citing, Cited by count]
- Related works, Concepts, SDGs, MeSH, ngrams
- Type, Paratext, Retraction status, Language
- Source [Name, Type, Publisher, OA, DOAJ]

**Author**
- Name (display, alternatives), ID, ORCID
- Works [Counts, Citations, H’index, i10_index, Concepts]
- Last known institution

**Source**
- Name
- ID, ISSN, MAG, Wikidata
- Hosting organizations, Societies
- OA? In DOAJ?
- APC price & currency
- Works [Counts, Citations, Concepts]

**Institution**
- Name [Display, Acronyms, Alternatives], ID, Type
- Location [City, Country, Region, Lat/Lon]
- Repositories, Associated institutions
- Roles (institution, publisher, funder)
- Works [Counts, Citations, Concepts]

**Concepts**
- Name
- ID, Wikidata
- Concept level, Ancestor concepts
- Description (multiple languages), Related concepts
- Works count, Cited by count
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

- Works: DOI
- Citations: Crossref
- Authors: ORCID
- Journals: ISSN
- Institutions: ROR
- Concepts: Wikidata
Surprised it’s free and better? See what others have to say.
"I am the main developer of VOSviewer, one of the most popular software tools worldwide for visualizing scientific literature based on bibliographic data. VOSviewer supports a large number of bibliographic databases. However, most of these databases have important limitations:

- They require an expensive subscription (Web of Science, Scopus).
- Their coverage of the scientific literature is limited (Web of Science, Scopus).
- They allow only small amounts of data to be exported (Web of Science, Scopus, Dimensions).
- They are restricted to specific disciplines (PubMed, Europe PMC).
- There are major gaps in their data (PubMed, Europe PMC, Crossref, OpenCitations), and/or
- Downloading data is very slow (Crossref, OpenCitations).

OpenAlex is of crucial importance for VOSviewer users because it offers a better performance than other databases on all the above-mentioned criteria.

VOSviewer users regularly run into problems because of the limitations of bibliographic databases. When they ask me for help, I often refer them to OpenAlex as an alternative database that is likely to offer a solution to their problem. I consider OpenAlex to be a fundamental building block for an ecosystem of open infrastructures for high-quality research analytics."

— Nees Jan van Eck, Centre for Science and Technology Studies (CWTS), Leiden University
Testimonials

"OpenAlex helps significantly optimize the performance of our literature search, and thus shorten our overall R&D time."
— Trang Le, Bristol Myers Squibb

"The OpenAlex API helps us to consolidate and accelerate our data collection."
— Michaela Voigt, Technische Universität Berlin (University library)

"We have now moved 100% to OpenAlex...data and metadata are very complete."
— André Vermeij, Kenedict

"OpenAlex is CO2 and we can share everything without worrying or talking to lawyers! It is really great."
— Tom Thiele, Max-Planck-Institute for demographic research

"OpenAlex is highly recommended for its fast, open, and high-quality scholarly data."
— Tim Wölfe, Local Citation Network

"OpenAlex gives me an open-source alternative that’s large, free, and has an easy-to-use API."
— Chris B., PhD student at US research university

"The best solution for academic use cases like ours that are working to further diversity and equity in research."
— Chinna Dankhra, Georgia Institute of Technology

"This is a powerful tool that has already cemented itself as an integral and vital part of bibliographic studies."
— Eric Scharles, Iowa State University Library

"OpenAlex is the best in terms of completeness, data quality, and ease of use...OpenAlex is better than the paid data sources."
— Adam Day, Clear Skies

More online at: openalex.org/testimonials
OpenAlex does have limitations.
OpenAlex limitations

- **Bias**: Inherits from sources, esp Crossref and MAG
- **Research base**: Not much, but rapidly growing
- **Coverage**: No patents. Limited software, datasets
- **Stability**: Data changing (improving) monthly
OpenAlex History

May 2021- Microsoft announced MAG sunsetting
Dec 2021- MAG discontinued
Jan 2022- OpenAlex beta launched
May 2022- User Group launched
August 2022- Full text search
December 2022- Customer support ticket system
March 2023- Premium offering launched
July 2023- Improved author disambiguation launched
How you can access OpenAlex.
<table>
<thead>
<tr>
<th>Object</th>
<th>Last Modified</th>
<th>Timestamp</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>data/</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>browse.html</td>
<td>a year ago</td>
<td>2022-05-18 10:09:34</td>
<td>32 KB</td>
</tr>
<tr>
<td>LICENSE.txt</td>
<td>7 months ago</td>
<td>2022-12-30 11:04:09</td>
<td>7 KB</td>
</tr>
<tr>
<td>README.txt</td>
<td>2 months ago</td>
<td>2023-06-06 06:11:51</td>
<td>129 Bytes</td>
</tr>
<tr>
<td>RELEASE_NOTES.txt</td>
<td>a month ago</td>
<td>2023-07-11 19:14:20</td>
<td>5 KB</td>
</tr>
</tbody>
</table>

Showing 1 to 5 of 5 entries

Search:
OpenAlex web interface- in Beta!

GUI for accessing specific subsets of data, building API calls, and exporting data

Walk-through coming soon!
Functions enabled by the content and metadata structure.
Subsetting & Analysing the Data

All works

Logic layer

Keywords or Filters
- Year
- Institution
- ..... Concept

Results

Subset

Intelligence
Find specific works

e.g., Open access journal publications mentioning kelp with an author from Canada that was funded by NSERC

All works

Logic layer

Keywords or Filters

- “kelp”
- OA
- Canada
- journal
- Funded by NSERC

Results
Common intelligence use cases

e.g., Open access journal publications mentioning kelp with an author from Canada that was funded by NSERC
OpenAlex API: filter

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

{
  "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 Prevention",
      "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 API: group

```json
- group_by: [
  - {
      key: "closed",
      key_display_name: "closed",
      count: 136100
    },
  - {
      key: "bronze",
      key_display_name: "bronze",
      count: 62526
    },
  - {
      key: "green",
      key_display_name: "green",
      count: 49484
    },
  - {
      key: "gold",
      key_display_name: "gold",
      count: 33073
    },
  - {
      key: "unknown",
      key_display_name: "unknown",
      count: 21913
    },
  - {
      key: "hybrid",
      key_display_name: "hybrid",
      count: 21913
    }
]```
Additional Resources

**OpenAlex Resources**
documentation: [https://docs.openalex.org/](https://docs.openalex.org/)
tutorial: [https://docs.openalex.org/quickstart-tutorial](https://docs.openalex.org/quickstart-tutorial)
google user group: [https://groups.google.com/g/openalex-users](https://groups.google.com/g/openalex-users)
help tickets: [https://openalex.org/help](https://openalex.org/help)
upcoming webinars: [https://openalex.org/webinars](https://openalex.org/webinars)

**Other Resources**
R Library to interface with OpenAlex APIs: [https://docs.ropensci.org/openalexR/](https://docs.ropensci.org/openalexR/)
OpenAlex is open and ready for use. Go play with it!

https://openalex.org/feedback
SKG content coverage sources

* 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
⅓ of works are in non-English language

source: our API!