

RR 59 - Appendix A: The Open Access community and its vocabulary

Instead of dealing with one for-profit company, the open access data curation model is decentralized among different non-profit organizations. Listed below are players participating in the CWTS Leiden OpenAlex ranking 2023.

OurResearch (<https://ourresearch.org/>) is the parent organization for open-source tools such as **OpenAlex** and **Unpaywall** (<https://unpaywall.org/>), launched in 2016 as a free source of scholarly articles. OurResearch also works with Clarivate, Elsevier, Dimensions, and others. For more information on the launch of Unpaywall, see Chawla, 2017, <https://doi.org/10.1038/nature.2017.21765>

Other feeder sites also include:

Directory of Open Access Journals (DOAJ) (<https://doaj.org/>), a curated list of open-access journals

Publishers provide metadata, including author affiliation, to **Crossref**, a membership organization <https://www.crossref.org/>, and **PubMed**, from the U S National Institute for Health, which has several tools for downloading the data, <https://pubmed.ncbi.nlm.nih.gov/download/>

Research Organization Registry (ROR) provides persistent IDs for research organizations. It is used to disambiguate institutions' names. It provides an ID that covers the parent organization and its affiliates and each affiliate also gets its ID (<https://ror.org/about/>). Webometrics used it for identifying universities in its January 2024 release.

COKI (Curtin Open Knowledge Initiative) is an initiative that provides an interactive dashboard with country-level data, [https://openknowledge.community/tracking-global-access-the-move-to-openalex-and-inclusion-of-2022-data/](https://openknowledge.community/Napier,K.,Neylon,C,&Diprose,J.(28June2023).Trackingglobalaccess-themovetoOPENAlexandinclusionof2022data.https://openknowledge.community/tracking-global-access-the-move-to-openalex-and-inclusion-of-2022-data/) .

Terminology:

METADATA– data that describes the data. <https://www.opendatasoft.com/en/blog/what-is-metadata-and-why-is-it-important-data/>

INGEST: Data ingestion refers to the tools & processes used to collect data from various sources and move it to a target site, either in batches or in real-time. <https://www.qlik.com/us/data-ingestion>

SCRAPING: see <https://techterms.com/definition/scraping>

OpenAlex technical documentation (Feb 2024); the document and the database are updated frequently (<https://docs.openalex.org/>).

BASIC SEARCHING OpenAlex interface, <https://openalex.org/>



Search and analyze the world's research.



Open access

Results for a search on open access

About 704,000 works

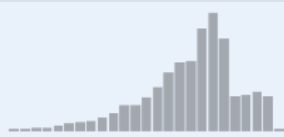
Relevance score ▾

- UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age
2015 · Cathie Sudlow, John Gallacher, et al. · *PLOS Medicine*
Cited by 6,609 PDF
- JASPAR: an open-access database for eukaryotic transcription factor binding profiles
2004 · Albin Sandelin · *Nucleic Acids Research*
Cited by 1,532 PDF
- Crystallography Open Database – an open-access collection of crystal structures
2009 · S. Gražulis, D. Chateigner, et al. · *Journal of Applied Crystallography*

Report



year



More...



open access



79.6%
560,600



institution



University of London

13,950



University of California System

13,530



French National Centre for Scientific Research

8,115



Chinese Academy of Sciences

6,865



University College London

6,603



Metrics for University of London

Alternate names (2)

Prifysgol Llundain

London University

Metrics

Citations	2,306,000
2yr mean_citedness	3.611
h index	455
i10 index	31,820