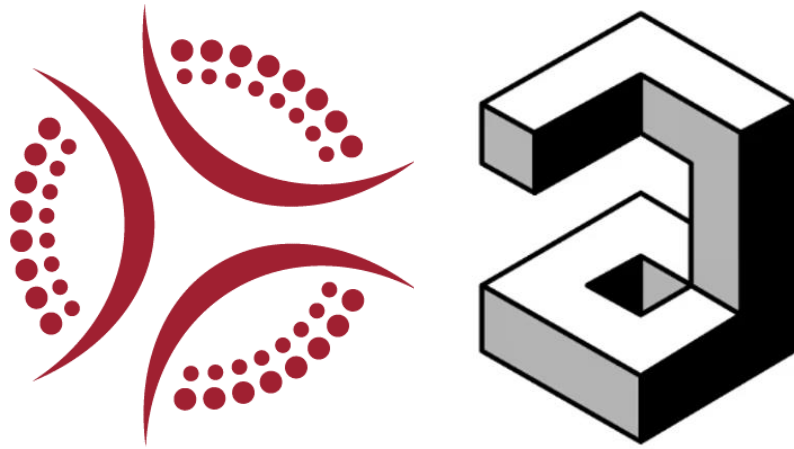


Hacky Morning – OpenAlex



KIM Workshop, 2022-05-03






Ablauf des Hackathons

- 09:30 Uhr Start mit Impulsvortrag
- 09:40 Uhr Gruppenfindung
- 10:10 Uhr Hacken
- 12:30 Uhr Zusammenkommen und kurze Auswertung
- 13:00 Uhr Ende

OpenAlex

- OpenAlex ist eine **frei verfügbare** Datenbank mit **offenen Metadaten** (CC0) zu wissenschaftlichen Arbeiten, Autor*innen, Institutionen u.a. und bietet **Zugriff über eine API** und **Datendumps** zum downloaden.
- OpenAlex wird von [OurResearch](#) seit ca. Mitte 2021 gepflegt und Daten kommen vom Microsoft Academic Graph, Crossref, institutionelle Repositorien (über OAI-PMH), ...
- Zugriff über API seit Anfang 2022

Entitäten in OpenAlex

-  **Works:** Artikel, Bücher, Datensätze etc.
-  **Authors:** Personen, welche ein Werk erstellen
-  **Venues:** Zeitschriften, Repositorien
-  **Institutions:** Universitäten und andere Einrichtungen
-  **Concepts:** fachliche Einsortierung der Werke

API-Zugriffe allgemein

- Basis-URL der API: `https://api.openalex.org/`
- Entity Endpoints: `/works`, `/authors`, `/venues`, `/institutions`, `/concepts`
- GET-Abfragen, z.B. direkt im Browser eingeben (ohne API-Key)
- Resultat in JSON
- Tages-Limit: 100.000 Abfragen; "polite pool" = E-Mail angeben
- Dokumentation: <https://docs.openalex.org/>

Erstes Beispiel für eine API-Abfragen

- Wie viele Werke sind aktuell in OpenAlex enthalten? - Alle Werke in OpenAlex

→ <https://api.openalex.org/works>

```
{
  meta: {
    count: 209337413,
    db_response_time_ms: 10,
    page: 1,
    per_page: 25
  },
  results:
  [...]
```

Weitere Beispiele mit Filter

- Alle Bücher:

→ <https://api.openalex.org/works?filter=type:book>

- Alle Werke von Autor*innen der Universität Mannheim:

→ <https://api.openalex.org/works?filter=institutions.id:I177802217>

- Alle Werke, welche "semantic web" im Titel enthalten (Phrasensuche):

→ <http://api.openalex.org/works?filter=title.search:semantic web>

- Welche interessanten Fragestellungen kann man durch solche Abfragen in OpenAlex direkt beantworten?



Komplexere Filtermöglichkeiten

- Alle Bücher von Autor*innen der Universität Mannheim (UND):

→ <https://api.openalex.org/works?filter=institutions.id:I177802217,type:book>



- Alle Autor*innen der Universität Mannheim ODER dem ZEW:

→ http://api.openalex.org/authors?filter=last_known_institution.id:I177802217|I177364143



- Alle Institutionen, welche nicht in den USA sitzen (NOT):

→ https://api.openalex.org/institutions?filter=country_code:!us



- Autoren von deutschen, schweizer oder österreichischen Institutionen mit mehr als 200.000 Zitationen:

→ /authors?filter=last_known_institution.country_code:DE|CH|AT,cited_by_count:>200000



Einzelne Einträge über IDs nachschlagen

- <https://api.openalex.org/works/W2741809807>
 - <https://api.openalex.org/works/doi:10.11588/ip.2020.2.73889> =
<https://api.openalex.org/works/https://doi.org/10.11588/ip.2020.2.73889>
 - <https://api.openalex.org/authors/https://orcid.org/0000-0002-6485-9434>
 - <https://api.openalex.org/institutions/https://ror.org/01n7gem85>
- Gibt es interessante Anreicherungen der eigenen Daten über einen Abgleich mit OpenAlex?



Weitere Optionen der API

- [Suche](#) im Titel + Beschreibung/Abstract
- [Sortieren](#)
- [Gruppieren und zählen](#) (z.B. OA-Verteilung bei Werken)
- [Autocomplete Endpoint](#)
- ...

Ideen für das Hacken mit OpenAlex



- Interessante Suchanfragen finden
- Anreicherungen z.B. auch Zitationen oder Abstracts
- Daten aus OpenAlex ziehen und visualisieren
- Analyse der Abdeckung der persönlichen Publikationen, der Publikationen der eigenen Einrichtung oder alle Publikationen einer bestimmten wohlbekanntes Zeitschrift analysieren
- Suchalerts oder OpenAlex auf bestimmte Änderungen monitoren
- gemeinsam das [API-Tutorial](#) durcharbeiten und [Tipp-des-Tages-Reihe](#)
- ...