



AALBORG UNIVERSITY
DENMARK

Aalborg Universitet

Bridging data gaps in a Research Information Management System with OpenAlex

Vidmar, Søren

DOI (link to publication from Publisher):
[10.5278/AAUOPENALEXCRIS](https://doi.org/10.5278/AAUOPENALEXCRIS)

Creative Commons License
CC BY 4.0

Publication date:
2024

Document Version
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

Citation for published version (APA):

Vidmar, S. (2024). *Bridging data gaps in a Research Information Management System with OpenAlex*. Poster presented at Officiel Lanceringskonference – Danmarks Forskningsportal, Frederiksberg, Denmark.
<https://doi.org/10.5278/AAUOPENALEXCRIS>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Bridging data gaps in a Research Information Management System with OpenAlex

Author

Søren Vidmar
sv@aub.aau.dk
0000-0003-3055-6053



Affiliation

Aalborg University Library



How can you discern the absence of something when you lack a comprehensive source of all necessary information?

Identifying data gaps within a research information management system presents a considerable challenge. It hinges on the available resources, licenses, organizational procedures, and of course time. Promising endeavors such as OpenAlex could offer us significant assistance.



Consequently, we have begun experimenting with new tools that could aid us in both detecting and bridging these gaps eventually.

CRISearch and OpenAlex2RIS are experimental tools that can help you identify and close the data gap in your Current Research Information System (CRIS)



1 Introduction

Data within a CRIS must be as complete as possible to provide a dependable base for distributing, showcasing and evaluating an institution's research contributions. Despite solid institutional processes and access to multiple systems to help us out, some records inevitably slip through the cracks each year. OpenAlex aids us in identifying missing records and potentially enhancing the metadata quality in our CRIS.

2 Objective

To initiate the creation of an easy-to-use tool that facilitates the identification and closure of data gaps within our CRIS system.

3 Identifying missing content with CRISearch - A PowerBI tool for the OpenAlex API

First step in closing the gap is identifying the publications missing in the CRIS.

The primary function of CRISearch is to compare the publications from our institution in OpenAlex against those in our CRIS, providing us with a comprehensive overview of any research outputs our institution may be lacking.

We use Elsevier's Pure as our CRIS, but the tool is system-agnostic as such, relying on DOI matching and a supplementary Lucene search for title matching to account for missing or faulty DOIs. By utilizing the APIs of OpenAlex and our CRIS, we can systematically identify gaps in our database, by narrowing down a list of publications which belong to our institution, but which we don't have in our CRIS.

| DOI | Publication Year | Possible match on title in CRIS? | Title | Type | Journal/Venue | Volume | Has volume? |
|---|------------------|----------------------------------|--|-----------------|--|--------|-------------|
| https://doi.org/10.1001/jama.network.open.2023.55716 | 2024 | No | Safety and Efficacy of Midline vs Peripherally Inserted Central Catheters Among Adults Receiving IV Therapy | journal-article | JAMA network open | 7 | Yes |
| https://doi.org/10.1002/9781394188789.ch1 | 2024 | Yes | The Necessity for Modernizing the Coupled Structure of Intelligent Transportation Systems and Multi-Energy Networks | other | | | No |
| https://doi.org/10.1002/9781394188789.ch6 | 2024 | Yes | Flexible Operation of Power-To-X Energy Systems in Transportation Networks | other | | | No |
| https://doi.org/10.1002/adfm.202313850 | 2024 | No | Cyanocarbonitrogen Based Low-Cost Polymer Donors for High Efficiency Organic Solar Cells | journal-article | Advanced Functional Materials | | No |
| https://doi.org/10.1002/advs.202304834 | 2024 | No | De Novo Atomistic Discovery of Disordered Mechanical Metamaterials by Machine Learning | journal-article | Advanced Science | | No |
| https://doi.org/10.1002/alz.13681 | 2024 | Yes | Mapping morbidity 10 years prior to a diagnosis of young onset Alzheimer's disease | journal-article | Alzheimer's & Dementia | | No |
| https://doi.org/10.1002/ctm2.11565 | 2024 | No | Pericardial delta like non-canonical NOTCH ligand 1 (Dlk1) augments fibrosis in the heart through epithelial to mesenchymal transition | journal-article | Clinical and translational medicine | 14 | Yes |
| https://doi.org/10.1002/dmrr.3775 | 2024 | No | The impact of sodium-glucose co-transporter-2 inhibitors on dementia and cardiovascular events in diabetic patients with atrial fibrillation | journal-article | Diabetes/Metabolism Research and Reviews | 40 | Yes |
| https://doi.org/10.1002/ehf2.14688 | 2024 | Yes | Computed tomography or chest X-ray to assess pulmonary congestion in dyspnoic patients with acute heart failure | journal-article | Esc Heart Failure | | No |
| https://doi.org/10.1002/jc.34851 | 2024 | Yes | Early mortality in children with cancer in Denmark and Sweden: The role of social | journal-article | International Journal of Cancer | | No |

4 Filling the gap with Python - OpenAlex2RIS

So identifying what's missing is one thing - filling the gap is another. To facilitate the integration of missing records, these scripts are capable of submitting either an individual DOI or a collection of DOIs in a CSV file to the OpenAlex API, producing RIS files in response. These RIS files can then be imported into our CRIS system.

```
Enter the DOI: 10.1016/S0021-9258(19)52451-6
Saved RIS file: output.ris
TY - JOUR
T1 - PROTEIN MEASUREMENT WITH THE FOLIN PHENOL REAGENT
J2 - Journal of Biological Chemistry
SN - 0021-9258
A1 - Oliver H. Lowry
A1 - N. J. Rosebrough
A1 - A. Farr
A1 - Rose J. Randall
PY - 1951
VL - 193
IS - 1
SP - 265
EP - 275
SO - https://doi.org/10.1016/S0021-9258(19)52451-6
LA - en
KW - folin phenol reagent
KW - protein
~ - https://doi.org/10.1016/S0021-9258(19)52451-6
AB - Since 1922 when Wu proposed the use of the Folin phenol reagent for the measurement of proteins (1), a number of modified analytical procedures utilizing this reagent have been reported for the determination of proteins in serum (2-6), in antigen-antibody precipitates (7-9), and in insulin (10).
```



5 Conclusion

Based on initial tests, it's advisable to approach publication import automation cautiously and ensure manual quality checks are in place. No database can encompass all sources, and while an open, comprehensive database by itself doesn't close the gap, it represents an important step in the right direction. Therefore, it's crucial that we explore OpenAlex more extensively in our work with information management. OpenAlex offers many opportunities we have yet to fully explore. Let's continue to delve into this resource and see what we can develop from it.

