OPEN METADATA SOURCES

COMPARING OPENAIRE TO CROSSREF

DATE: 07 MAY 2023

[PRELIMINARY VERSION]

Executive Summary

In this project, we assess and compare the value added by OpenAIRE to Crossref metadata, both in coverage of publications and other research output (with and without DOIs) as well as in coverage of metadata (including identifiers) for authors, institutions, publication venues and disciplines.

The report currently contains all the graphs comparing metadata coverage of OpenAIRE compared to Crossref, and of DOIs vs non-DOIs in openaire. More explanatory text, tables and interpretation of findings will be added in a later version.

Complete data and code are available on Github:
https://github.com/Curtin-Open-Knowledge-Initiative/open-metadata-report
All images and data belonging to this report are located in the directory
reports\run_20230507_crossref_openaire_1 in this repository.

Introduction and Background

In January 2022, OpenAlex was launched as a source of open bibliographic metadata. Intended both as a replacement of and improvement on Microsoft Academic, it provides structured data on publications, authors, institutions and publication venues.

Many tools, projects and services relied on Microsoft Academic as source of largely open metadata, and might consider switching to OpenAlex. More broadly, the launch of OpenAlex has increased interest in the potential of open metadata to enable discovery, linking and integration of data on research processes and outputs.

Unlike metadata from closed sources, open metadata can be combined and enriched to provide a rich open metadata landscape. Transparency and provenance allow identifying and addressing existing gaps and biases in coverage and quality.

In this project, we assess and compare the value added by OpenAIRE to Crossref metadata, both in coverage of publications and other research output (with and without DOIs) as well as in coverage of metadata (including identifiers) for authors, institutions, publication venues and disciplines.

Data sources

This report was run using the following tables as source data:

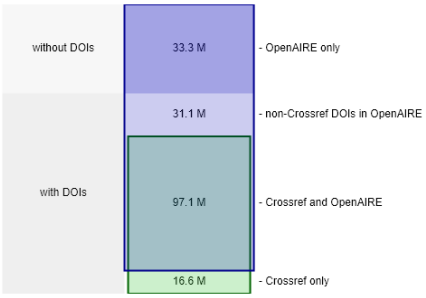
- Crossref: academic-observatory.crossref.crossref_metadata20221207
- OpenAIRE: academic-observatory.openaire.publication20221230

Complete data and code are available on Github:
https://github.com/Curtin-Open-Knowledge-Initiative/open-metadata-report
All images and data belonging to this report are located in the directory
reports\run_20230507_crossref_openaire_1 in this repository.

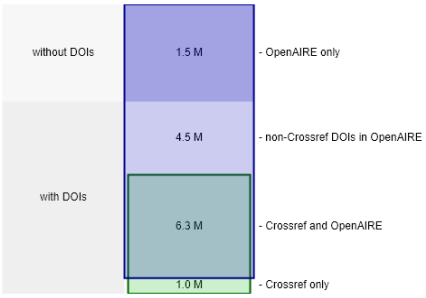
Coverage of OpenAIRE vs Crossref

Comparing coverage

Overview

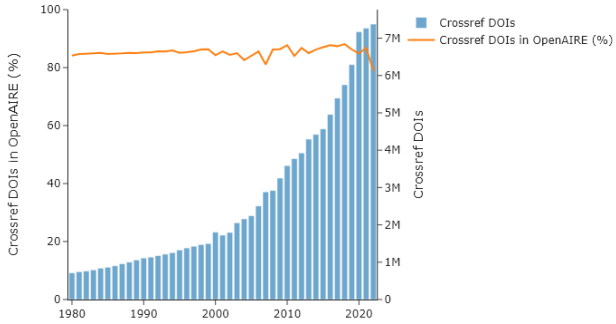


overall comparison - all time

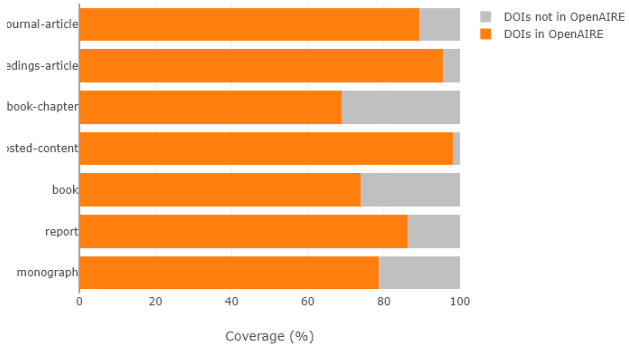


overall comparison - 2021

By year and publication type



coverage by publication date - all time

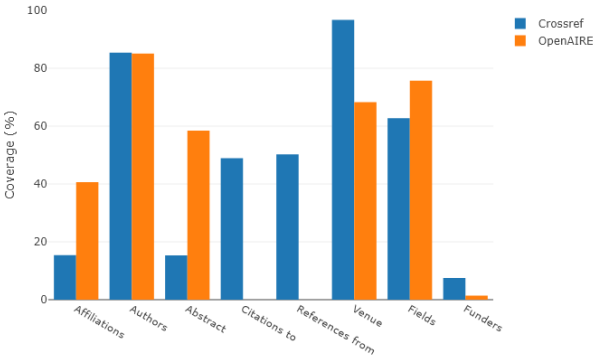


coverage by publication type - all time

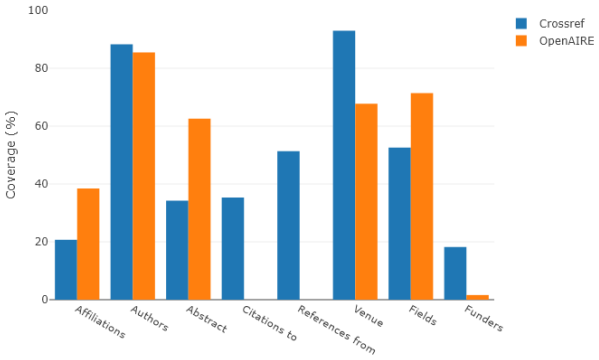
Value Add of OpenAIRE to Crossref

Overview

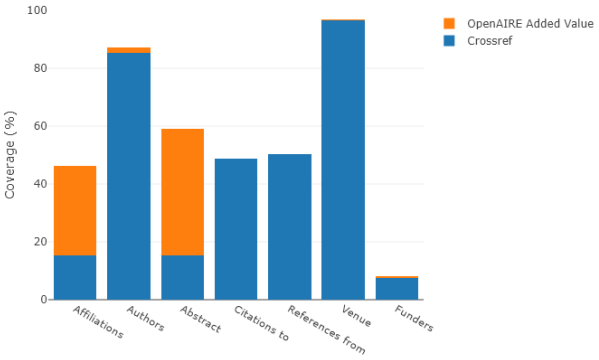
Comparing coverage of metadata types in Crossref and OpenAIRE



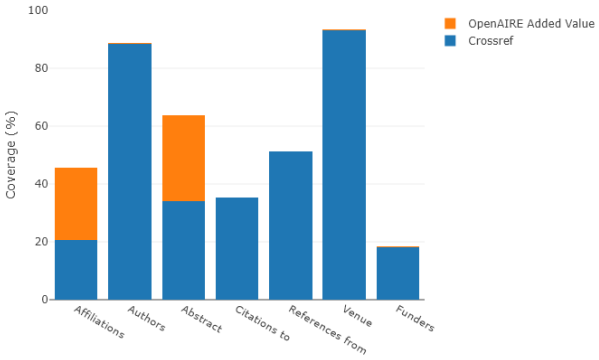
coverage comparison - all time



coverage comparison - 2021



coverage added value - all time

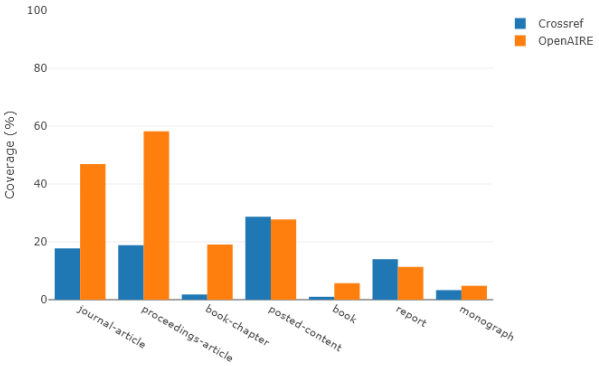


coverage added value - 2021

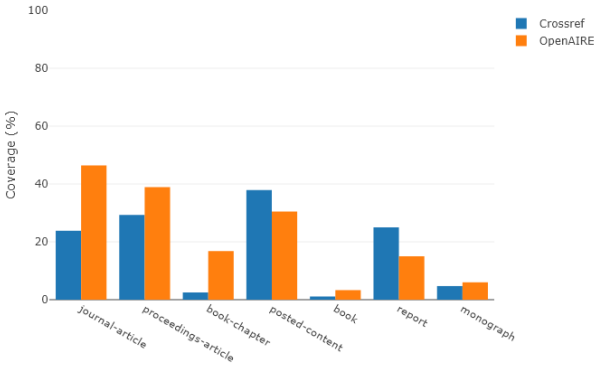
Details

Metadata coverage in OpenAIRE and Crossref by publication type

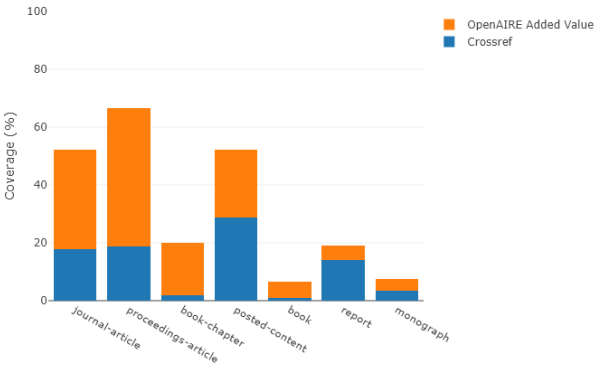
Affiliations



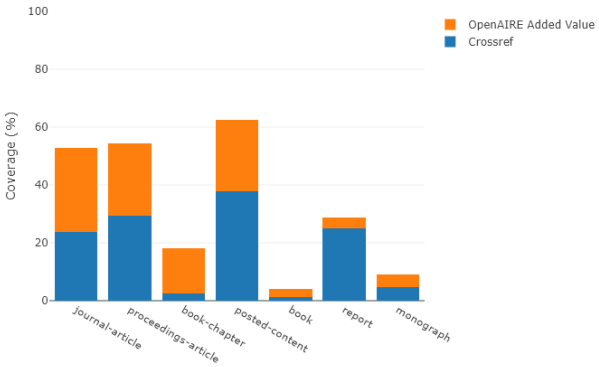
coverage comparison - all time



coverage comparison - 2021

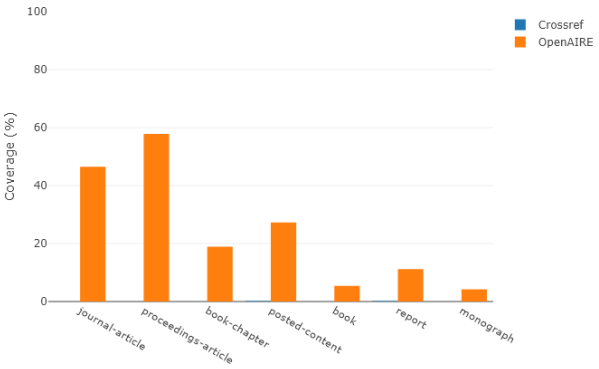


coverage added value - all time

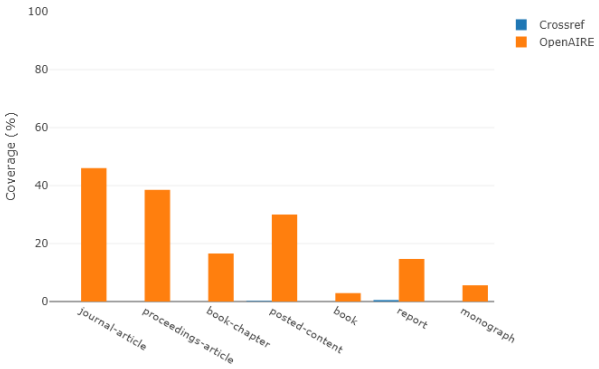


coverage added value - 2021

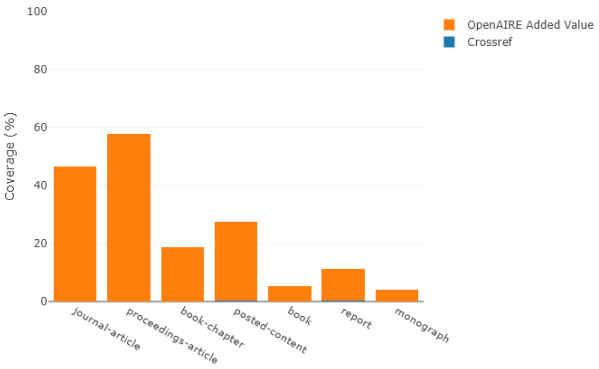
Affiliation RORs



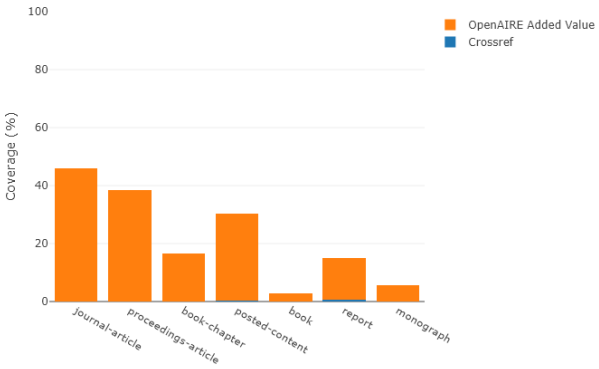
coverage comparison - all time



coverage comparison - 2021

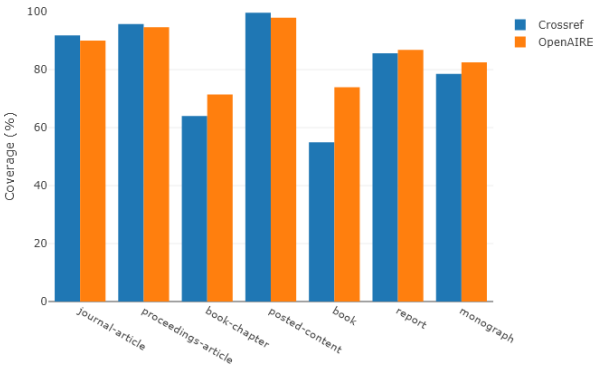


coverage added value - all time

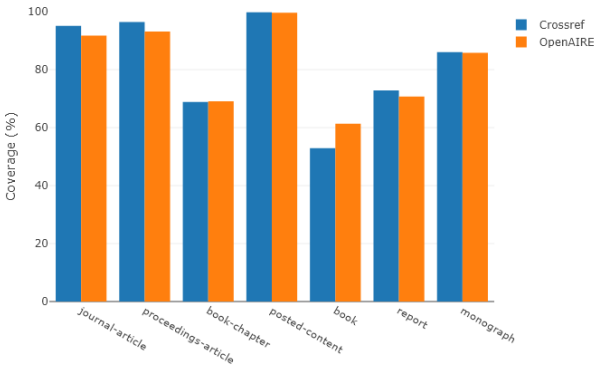


coverage added value - 2021

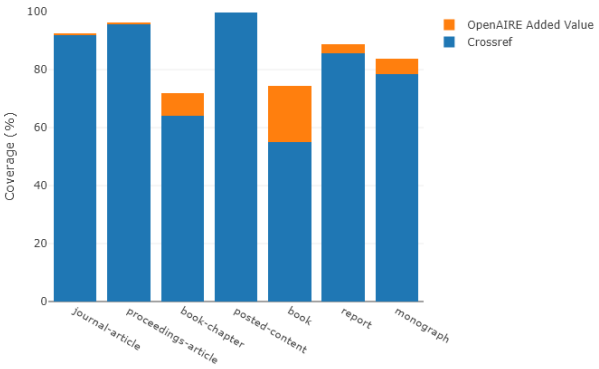
Authors



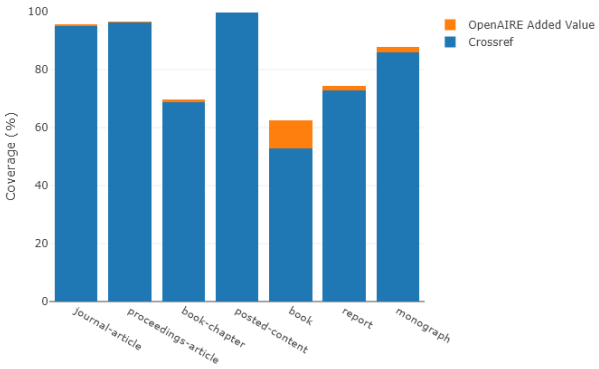
coverage comparison - all time



coverage comparison - 2021

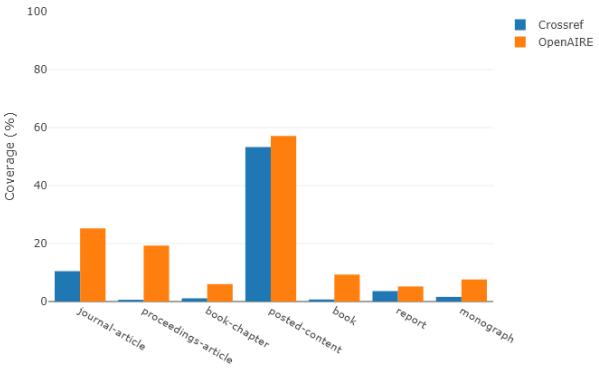


coverage added value - all time

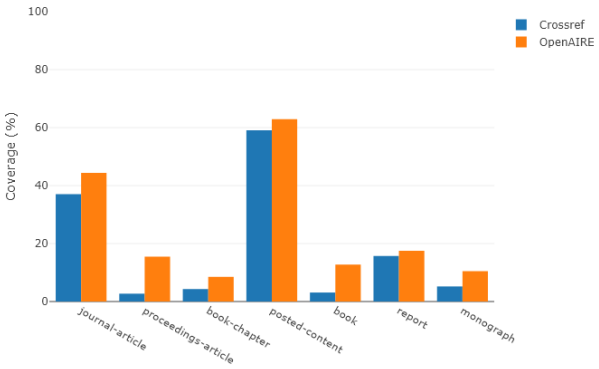


coverage added value - 2021

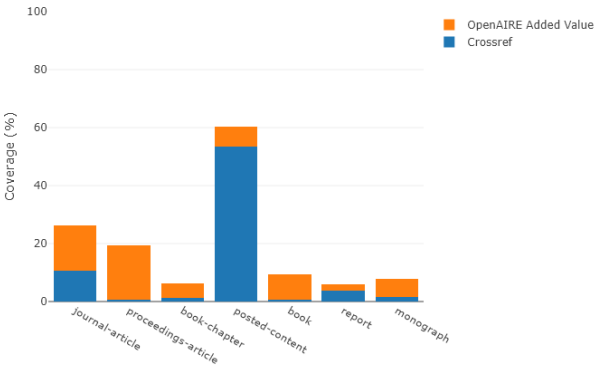
Author ORCIDs



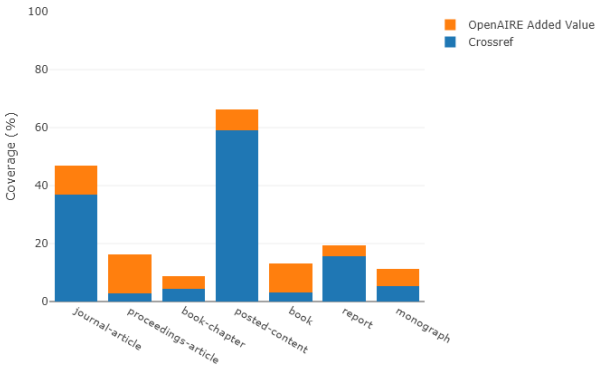
coverage comparison - all time



coverage comparison - 2021

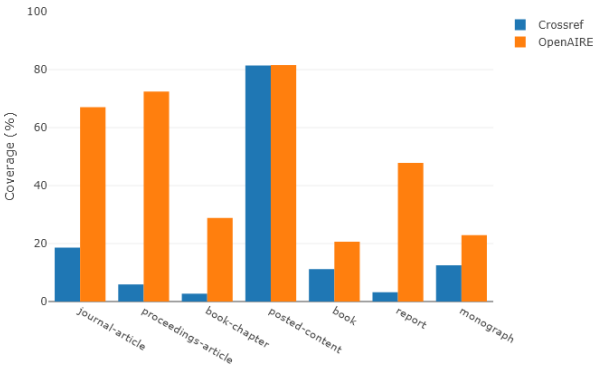


coverage added value - all time

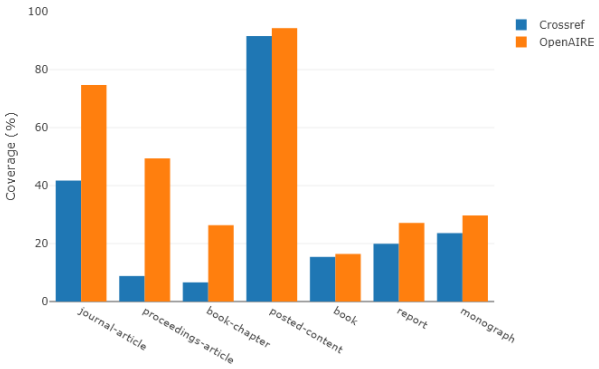


coverage added value - 2021

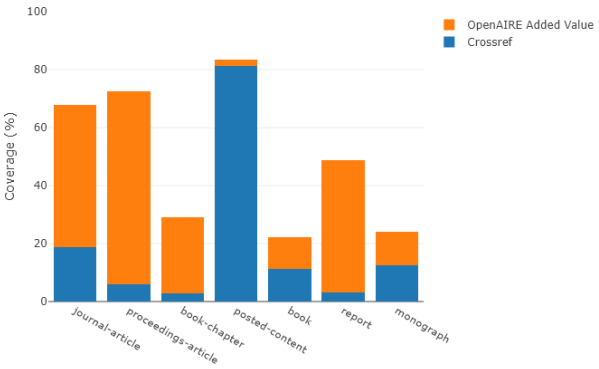
Abstract



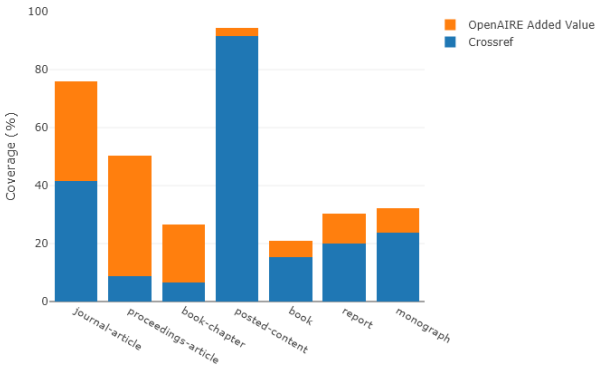
coverage comparison - all time



coverage comparison - 2021

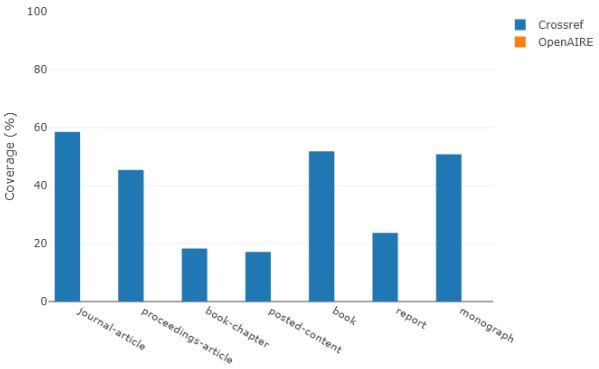


coverage added value - all time

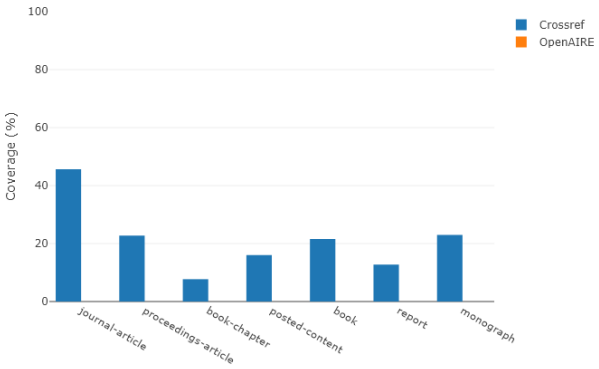


coverage added value - 2021

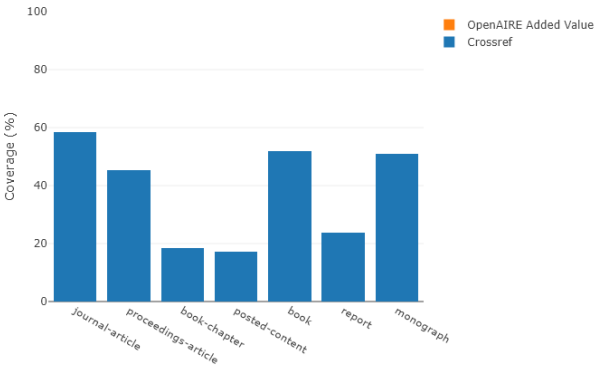
Citations to



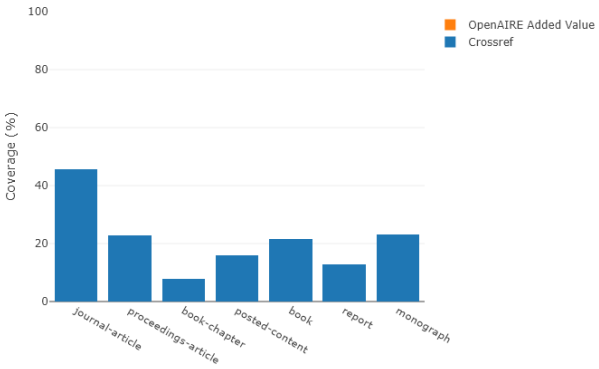
coverage comparison - all time



coverage comparison - 2021

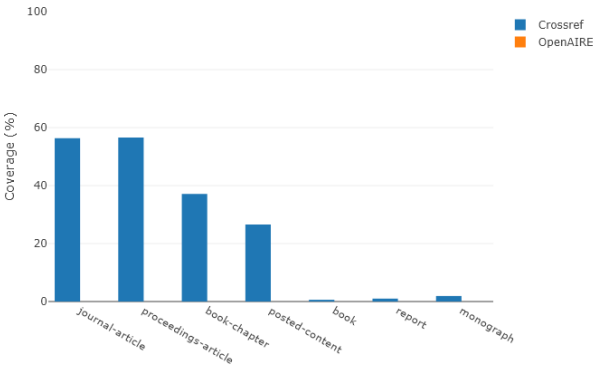


coverage added value - all time

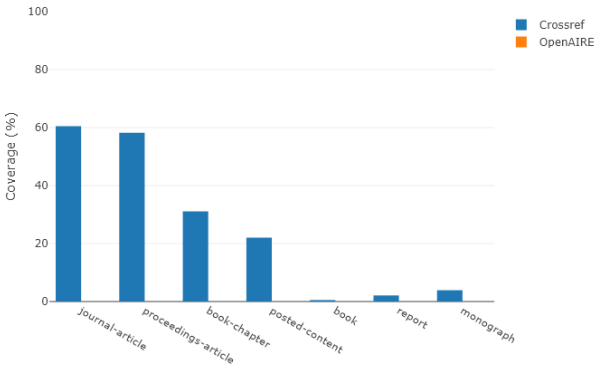


coverage added value - 2021

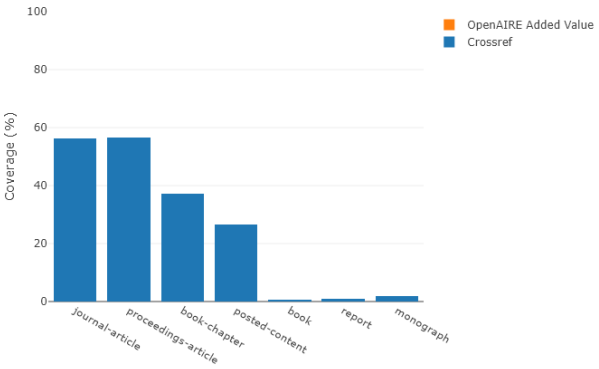
References from



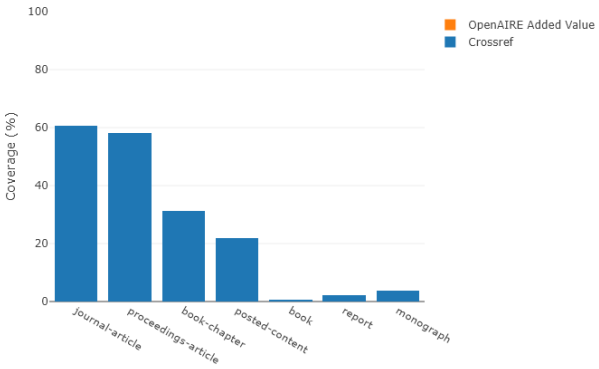
coverage comparison - all time



coverage comparison - 2021

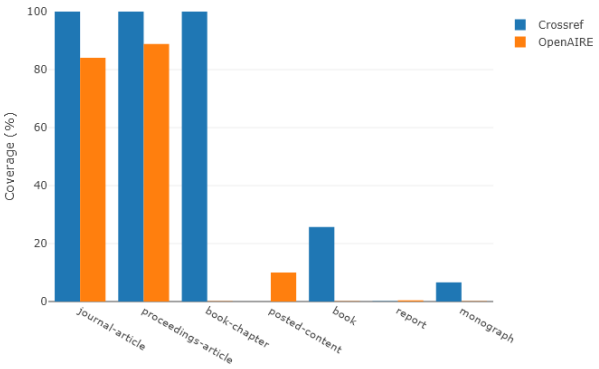


coverage added value - all time

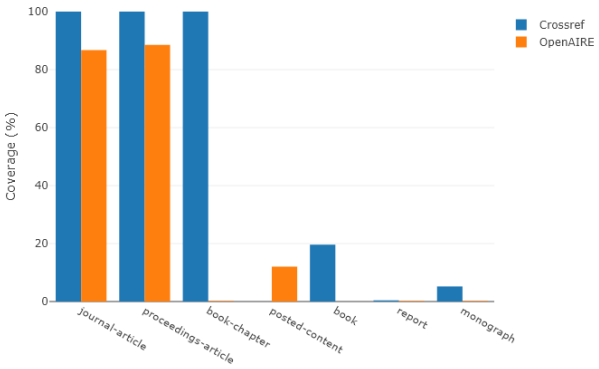


coverage added value - 2021

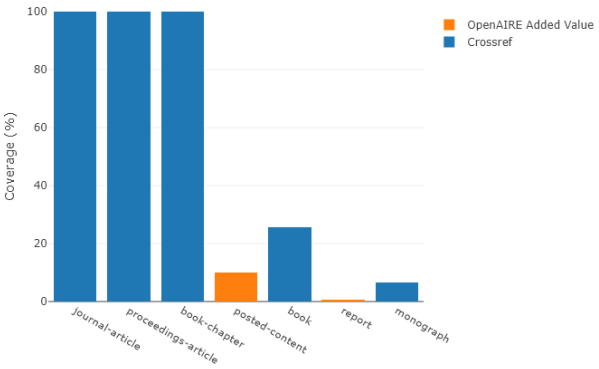
Venue



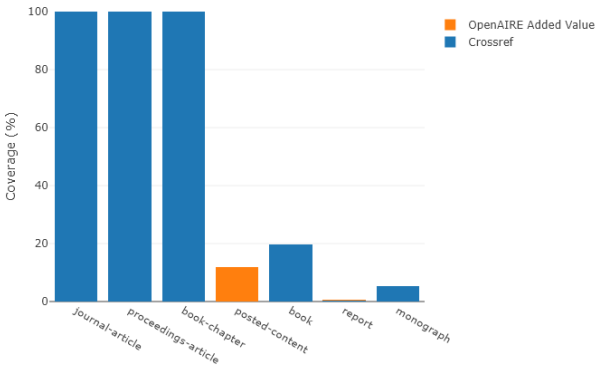
coverage comparison - all time



coverage comparison - 2021

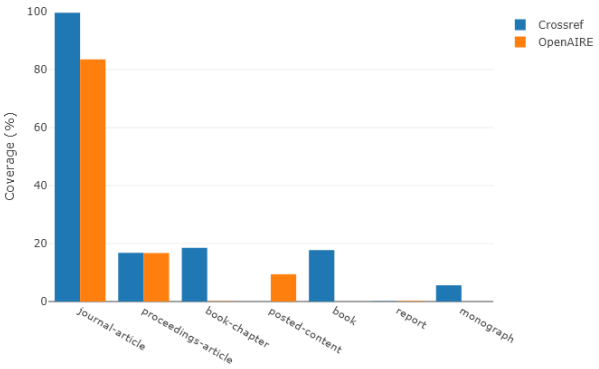


coverage added value - all time

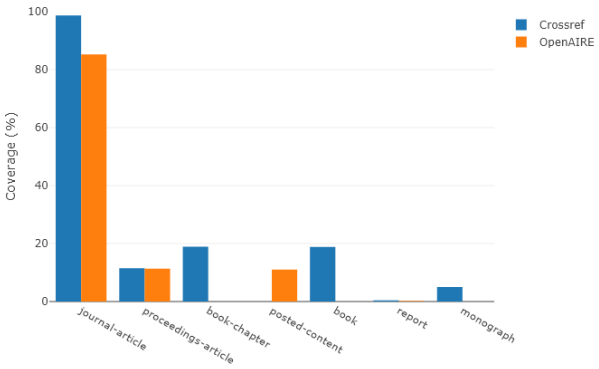


coverage added value - 2021

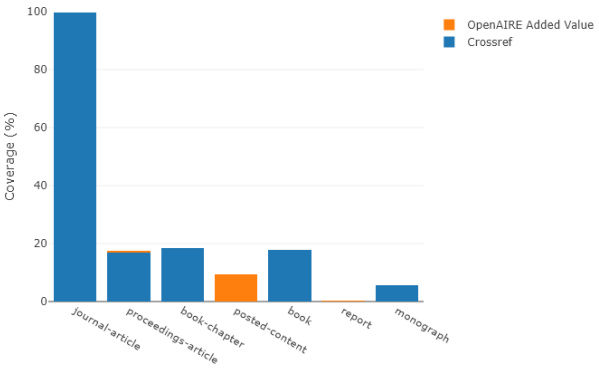
Venue ISSN



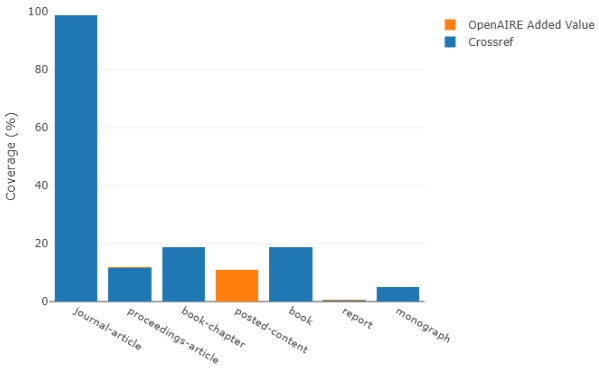
coverage comparison - all time



coverage comparison - 2021

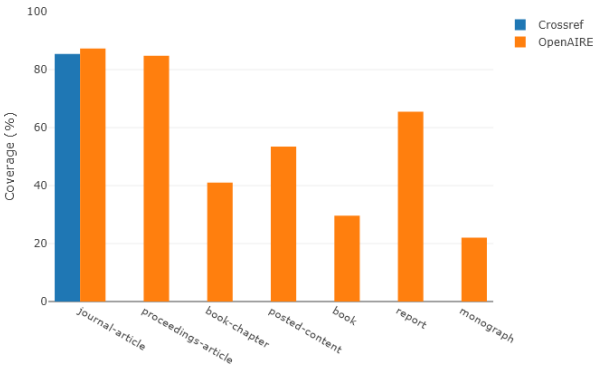


coverage added value - all time

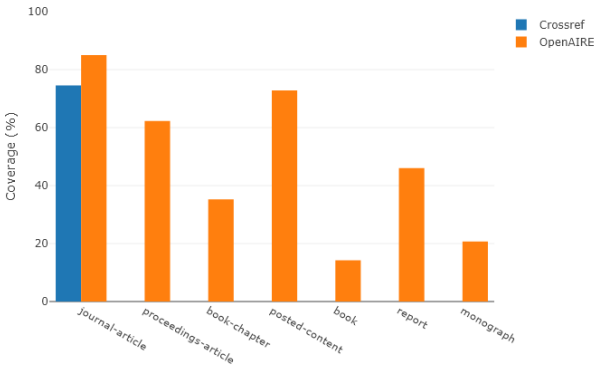


coverage added value - 2021

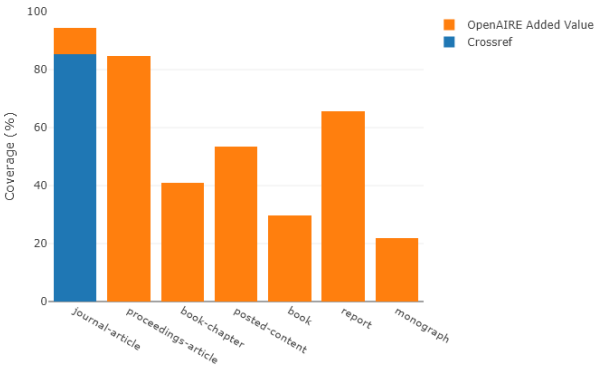
Fields



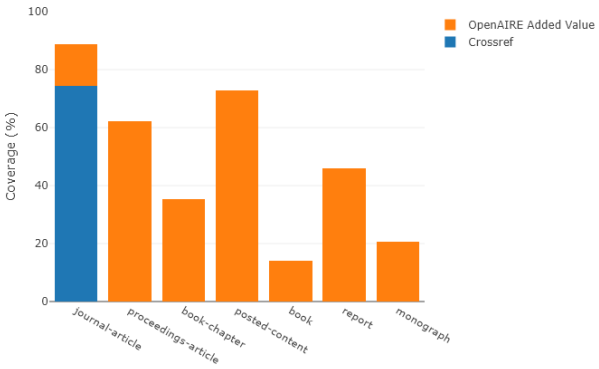
coverage comparison - all time



coverage comparison - 2021

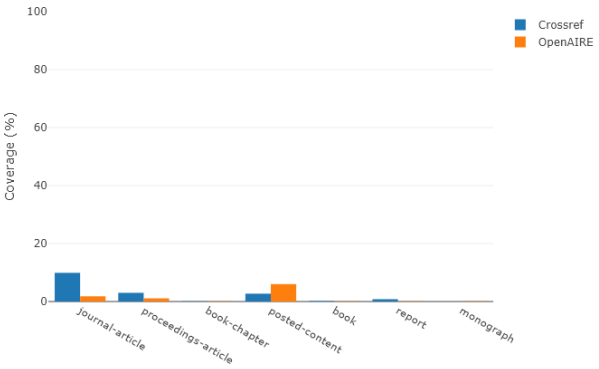


coverage added value - all time

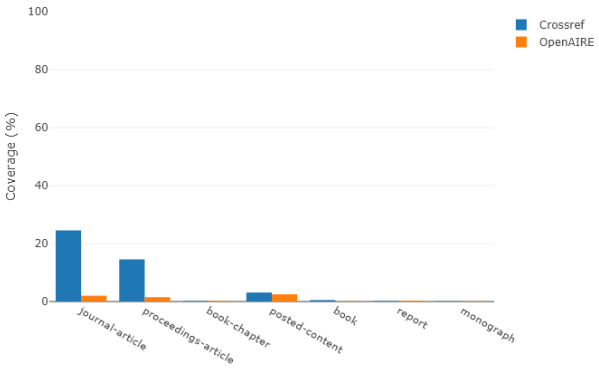


coverage added value - 2021

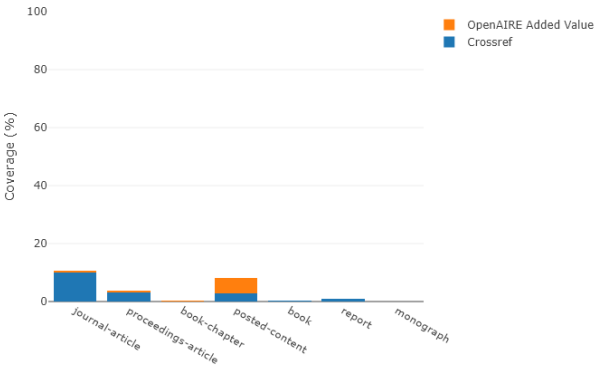
Funders



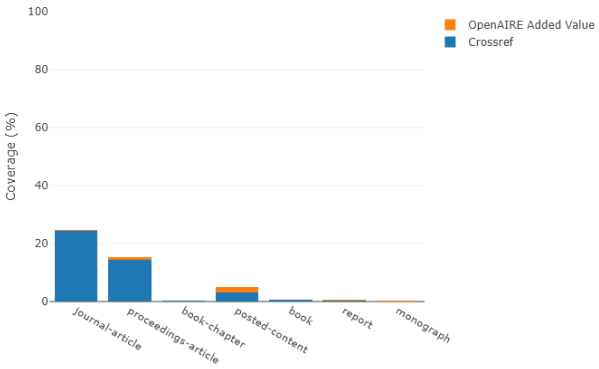
coverage comparison - all time



coverage comparison - 2021

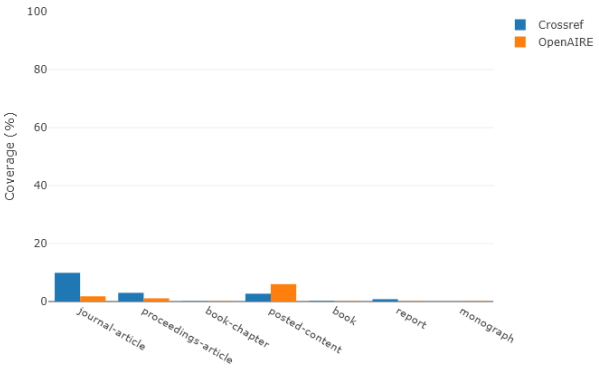


coverage added value - all time

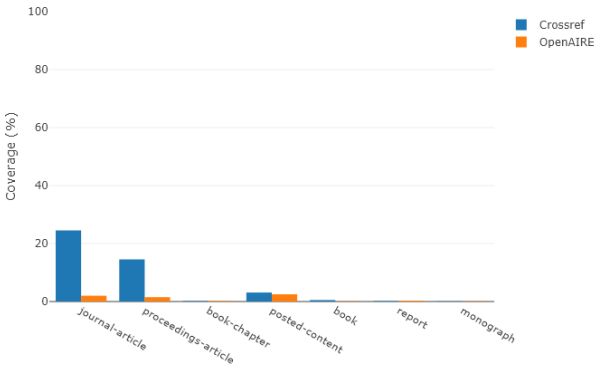


coverage added value - 2021

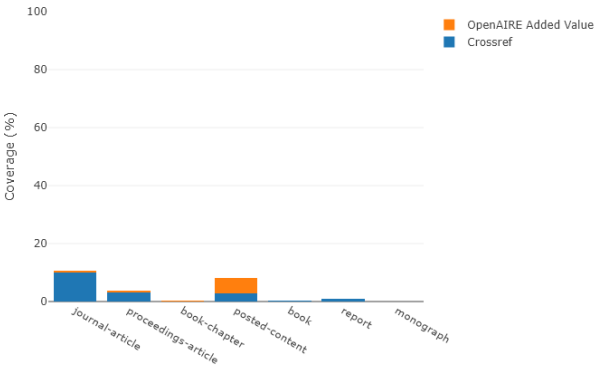
Funder Strings



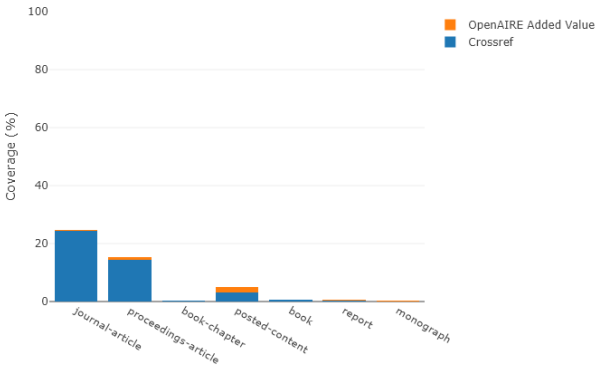
coverage comparison - all time



coverage comparison - 2021

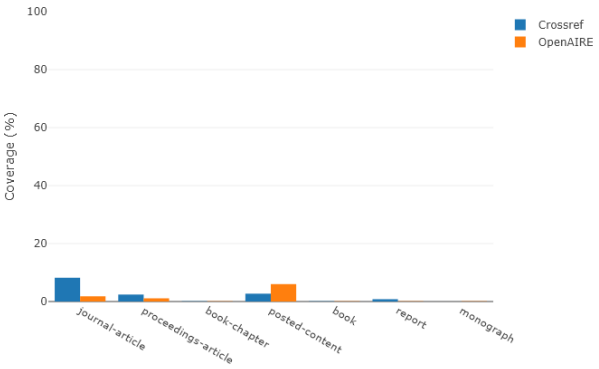


coverage added value - all time

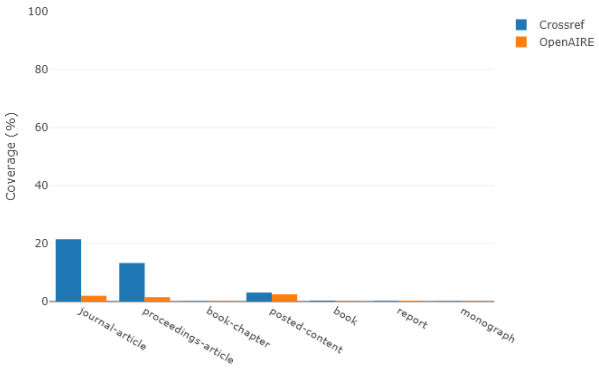


coverage added value - 2021

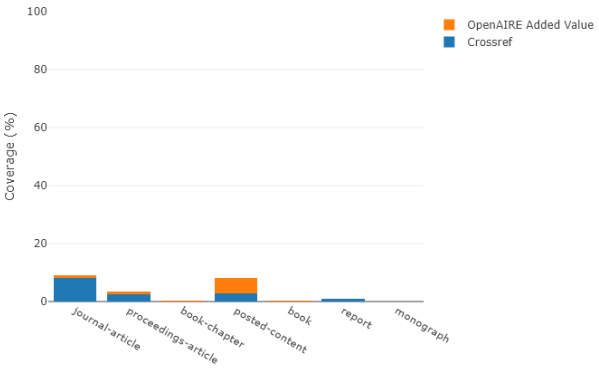
Funder Source IDs



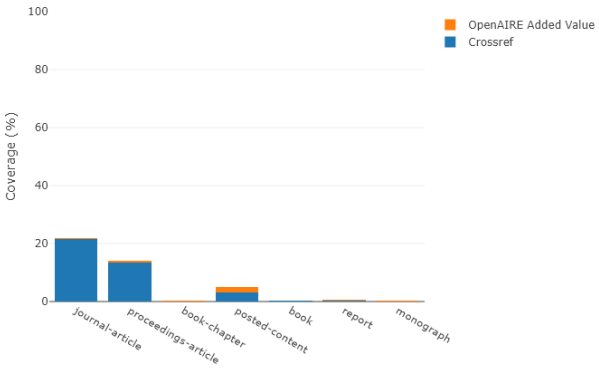
coverage comparison - all time



coverage comparison - 2021



coverage added value - all time

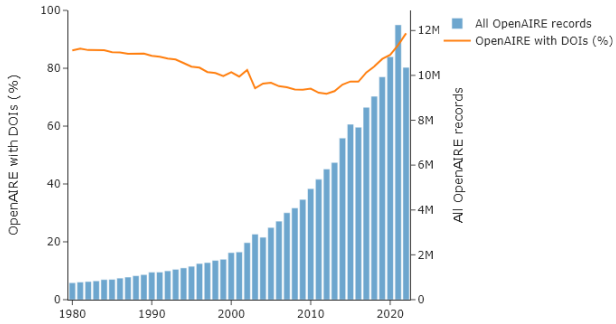


coverage added value - 2021

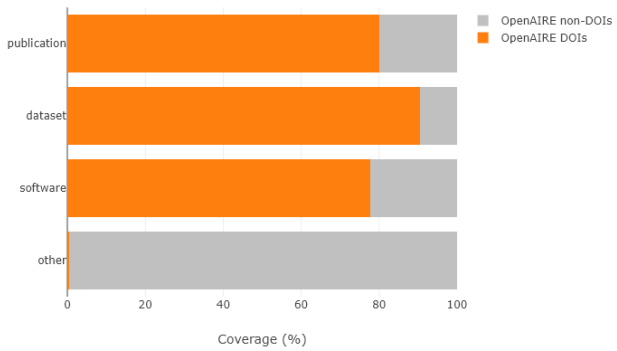
OpenAIRE Coverage Beyond DOIs

DOIs vs non-DOIs

By year and publication type



coverage by publication date - all time

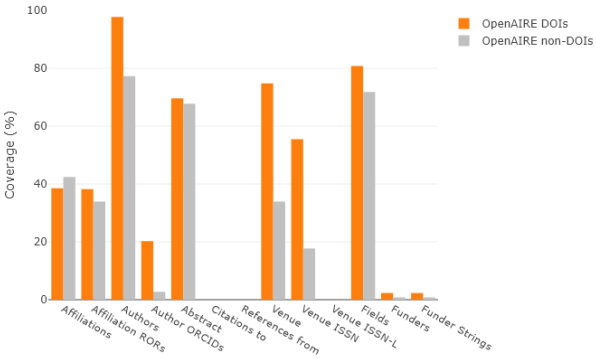


coverage by publication type - all time

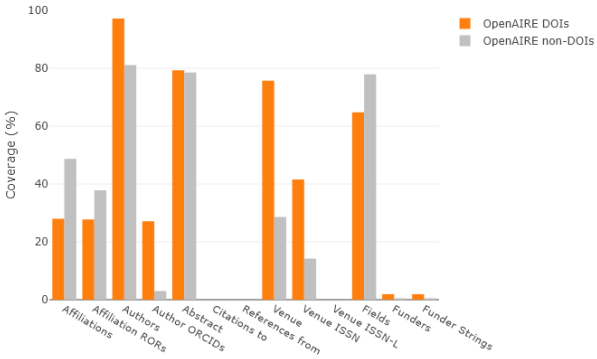
Metadata Coverage

Overview

Comparing coverage of metadata types for DOIs and non-DOIs in OpenAIRE



coverage comparison - all time

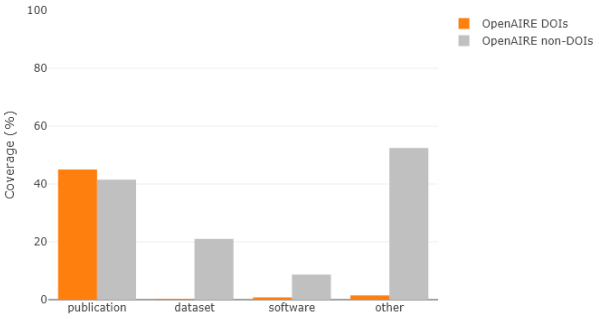


coverage comparison - 2021

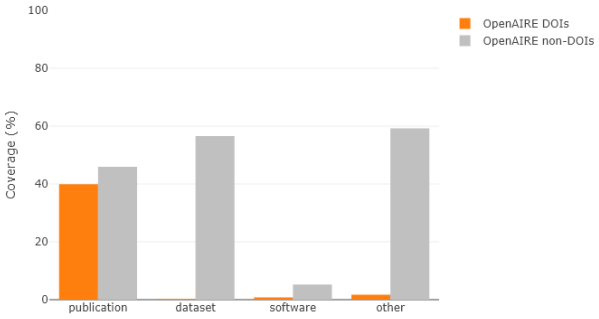
Details

Metadata coverage for DOIs and non-DOIs by publication type

Affiliations

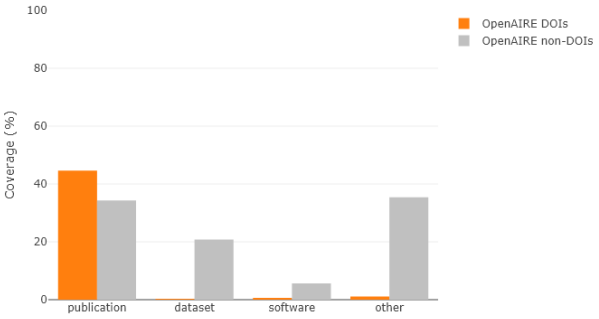


coverage comparison - all time

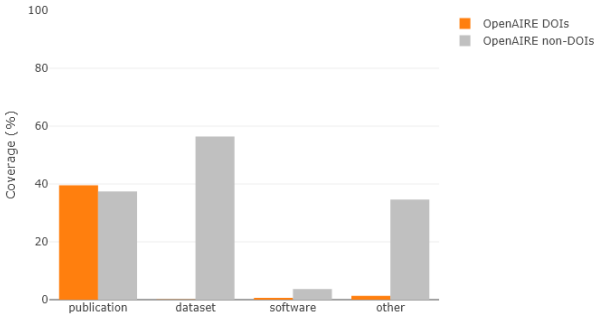


coverage comparison - 2021

Affiliation RORs

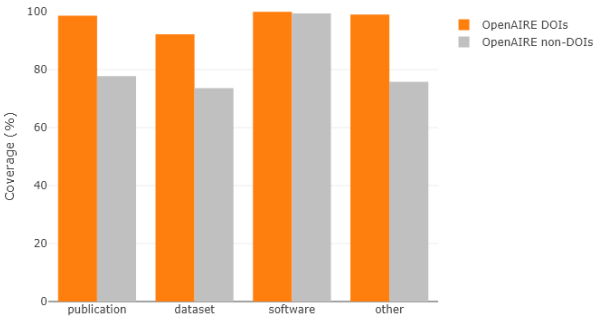


coverage comparison - all time

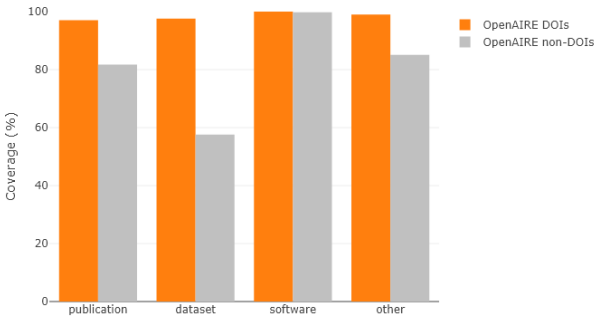


coverage comparison - 2021

Authors

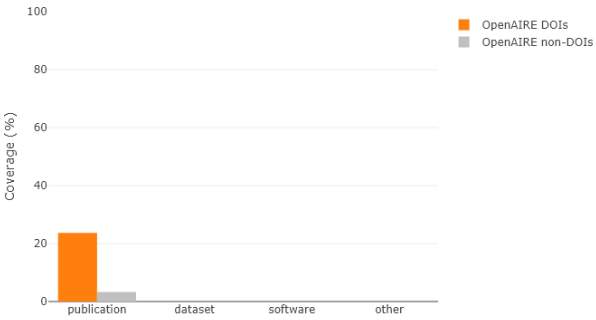


coverage comparison - all time

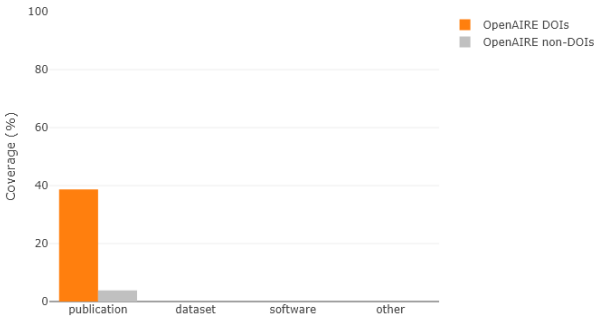


coverage comparison - 2021

Author ORCIDs

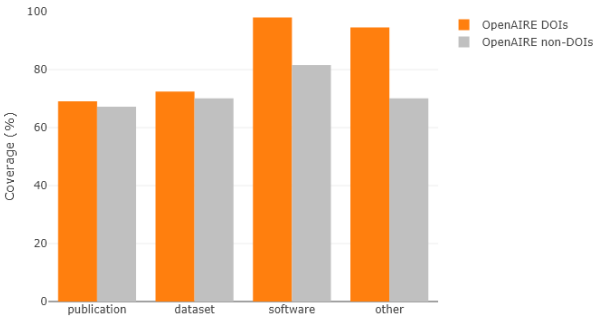


coverage comparison - all time

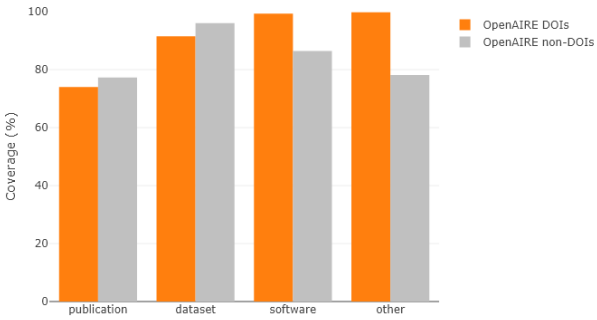


coverage comparison - 2021

Abstract

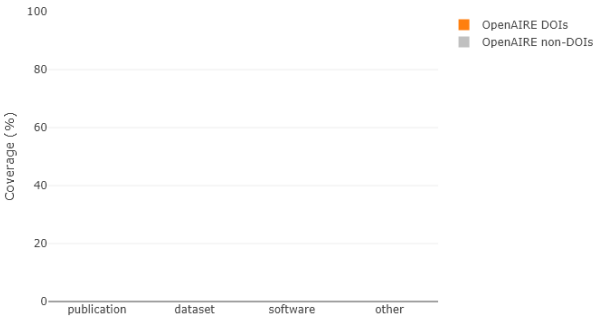


coverage comparison - all time

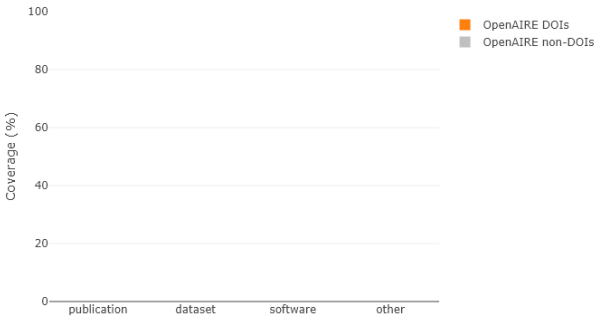


coverage comparison - 2021

Citations to

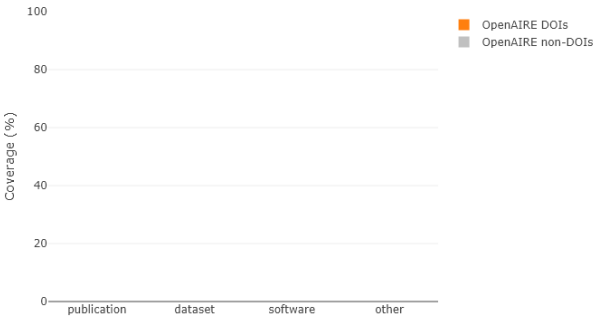


coverage comparison - all time

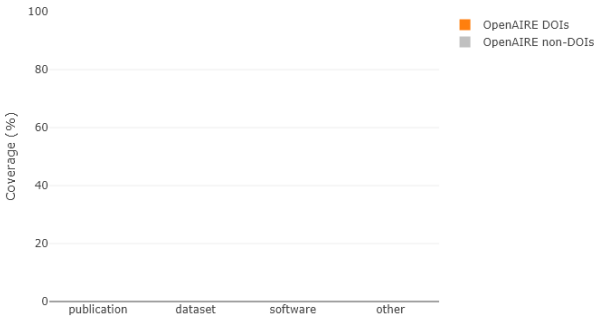


coverage comparison - 2021

References from

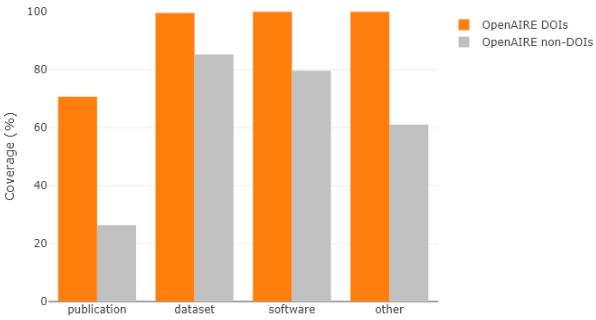


coverage comparison - all time

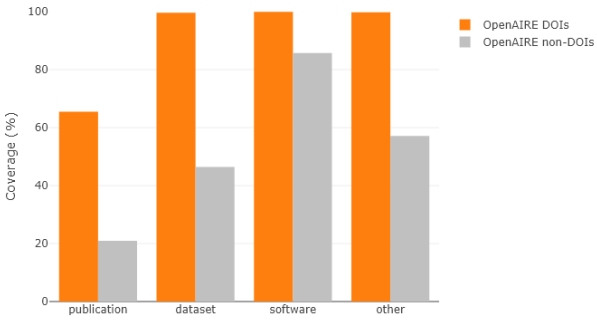


coverage comparison - 2021

Venue

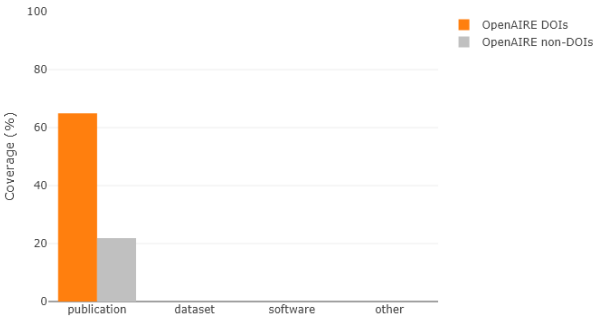


coverage comparison - all time

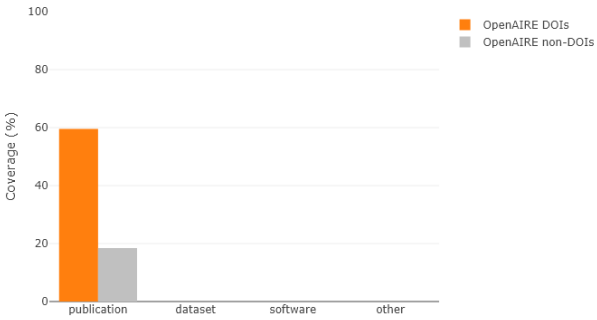


coverage comparison - 2021

Venue ISSN

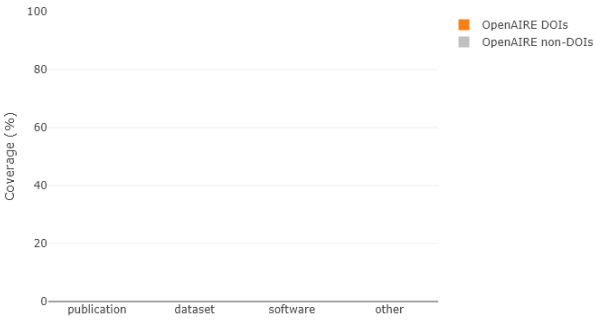


coverage comparison - all time

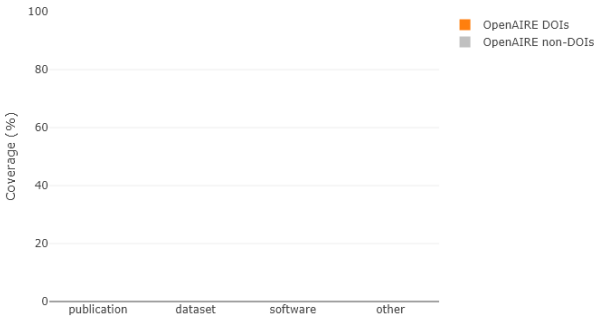


coverage comparison - 2021

Venue ISSN-L

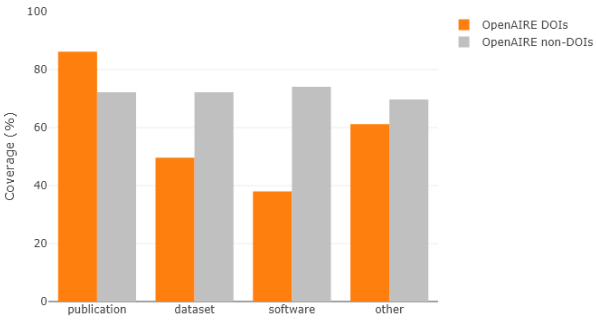


coverage comparison - all time

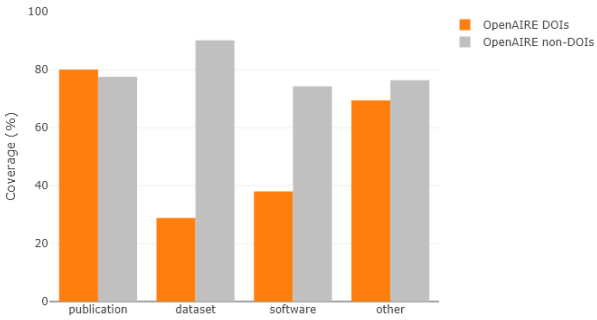


coverage comparison - 2021

Fields

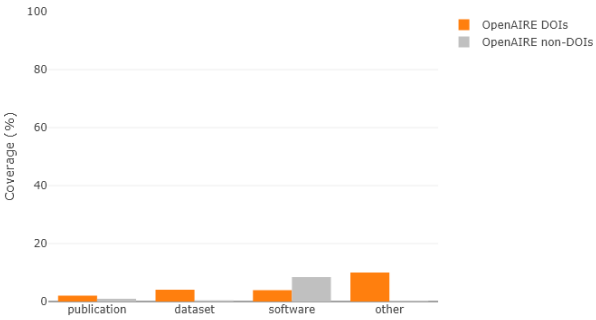


coverage comparison - all time

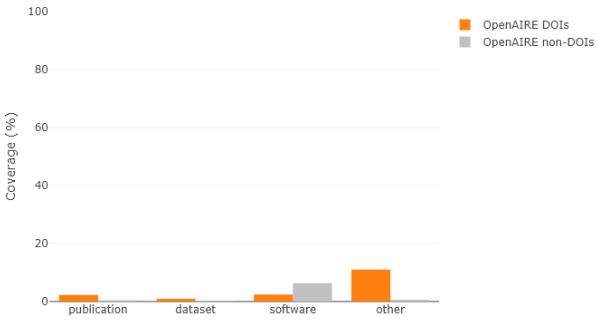


coverage comparison - 2021

Funders

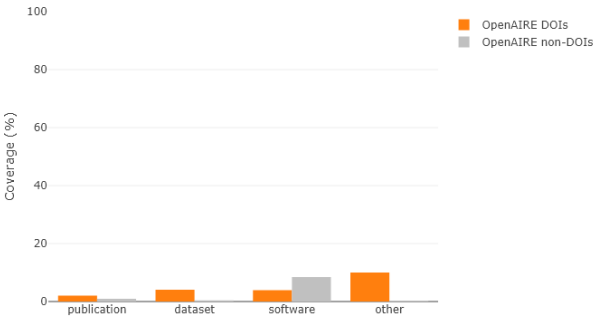


coverage comparison - all time

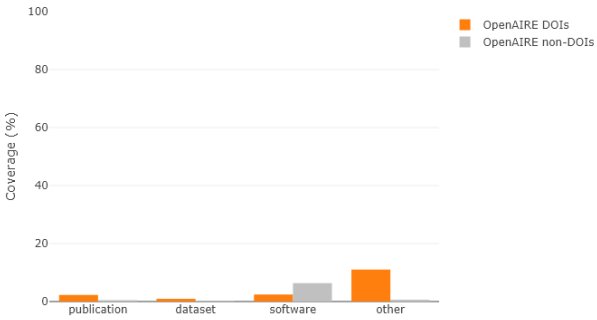


coverage comparison - 2021

Funder Strings



coverage comparison - all time



coverage comparison - 2021