

System Development: Data Enrichment Software Records with Citation Capture Pipeline

Taylor Jacovich and the ADS Team

ADS Users Group Meeting, 9-10 Nov. 2022



Software Records with Citation Capture Pipeline

Overview

- CitationCapture Pipeline handles ingest for certain software records

- Registers Zenodo records with ADS

Two types of records

Version: point to a specific version of the software

`property:release`

Software Source: refer to the software as a whole

- Forwards Zenodo records and Github urls to the Asclepias project

Asclepias: Collaboration between AAS, Zenodo, and ADS to create a tool for scientific software discovery

- Only ingests from reference lists of other ADS records

We do not ingest every Zenodo software record

Currently have ~12700 records

- Curators can non-destructively modify any metadata

Curation persists across metadata updates

Citation Capture Records

Associated Works

- All software records are automatically matched with any existing versions in ADS
 - The version name will reflect whatever versioning scheme the maintainers use

The image displays three panels of citation capture records, each with a dark background and white text. Each panel is titled 'FULL TEXT SOURCES' and includes a 'My Institution' and 'Publisher' section. Below these, there is a 'RELATED MATERIALS' section with a count in parentheses. The first panel shows 14 related materials, the second shows 8, and the third shows 3. Each panel also has a 'Show Less' button at the bottom.

Panel 1 (Left): FULL TEXT SOURCES. My Institution, Publisher. RELATED MATERIALS (14). Software Source. Version 0.9.11, Version 0.9.12, Version 0.9.13, Version 0.9.14, Version 0.9.15, Version 0.9.3, Version 0.9.5, Version 0.9.6rc1, Version 0.9.7, Version 1.0.0, Version 1.0.1, Version 1.0.2, Version 1.0.3. Show Less.

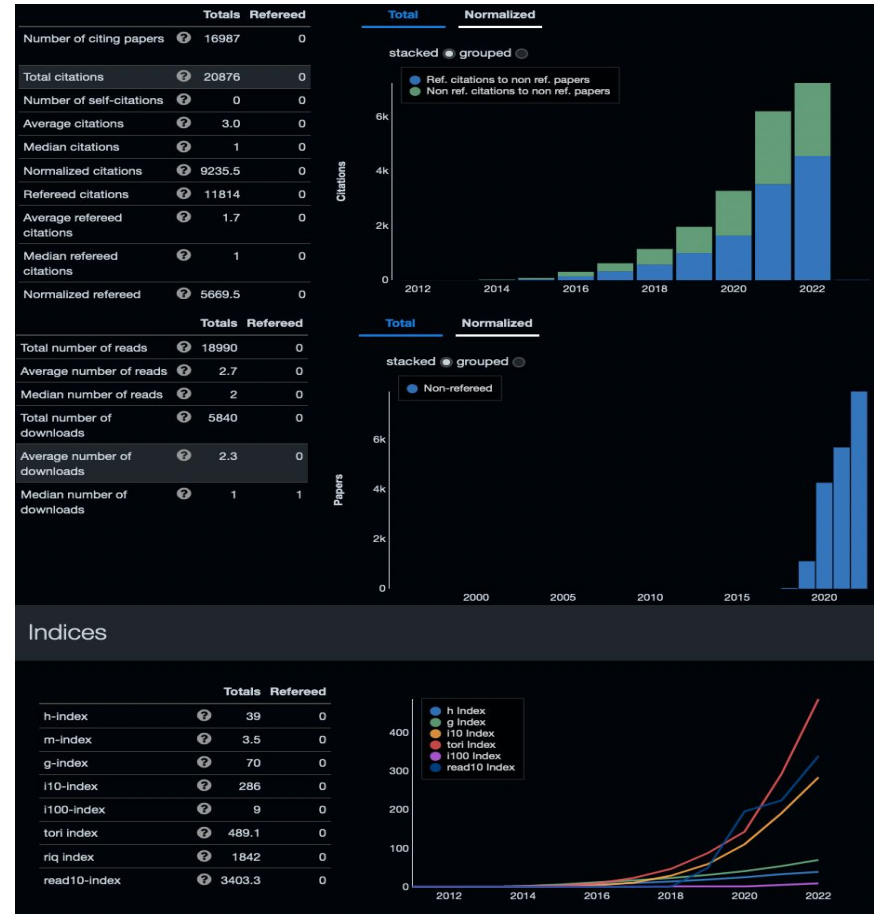
Panel 2 (Middle): FULL TEXT SOURCES. My Institution, Publisher. RELATED MATERIALS (8). Version 20.12.1, Version 20.3.1, Version 20160129, Version 20180822, Version 20190503, Version 20190830, Version 21.4.1, Version MAC OS 21.2.1. Show Less.

Panel 3 (Right): FULL TEXT SOURCES. My Institution, Publisher. RELATED MATERIALS (3). Version 3.1.1, Version FWSW-SDC-v2, Version parSDC_v1.

Citation Capture Records

Metrics

- Aggregate metrics can now include CitationCapture records
- Metrics can now be generated for individual CitationCapture records



Citation Capture Records

Potential Extensions

- Search function to aggregate associated works
- Better handling of Software Source records.