

# OpenCitations

**OpenCitations, an open e-infrastructure to foster maximum reuse of citation data**  
**IDCC22**

Ivan Heibi

Research Centre for Open Scholarly Metadata & Digital Humanities Advanced Research Centre (DHARC),  
Dipartimento di Filologia Classica e Italianistica, Università di Bologna, Bologna, Italia  
[ivan.heibi2@unibo.it](mailto:ivan.heibi2@unibo.it) – <https://www.unibo.it/sitoweb/ivan.heibi2/en>

Responsible for the Technical Infrastructure of OpenCitations  
[tech@opencitations.net](mailto:tech@opencitations.net) – [@opencitations](https://www.opencitations.net) – <http://opencitations.net>



# What is OpenCitations?

OpenCitations (<http://opencitations.net>) is an independent infrastructure organization dedicated to open scholarship and the publication of open bibliographic and citation data by the use of Semantic Web technologies

It provides:

- a data model: the OpenCitations Data Model (based on the SPAR Ontologies)
- bibliographic and citation data (CC0): OpenCitations Corpus, COCI, CROCI, CCC
- software: in our GitHub repository, released with open source licenses
- online services: REST APIs, SPARQL endpoints, dumps and interfaces

Our values:

- UNESCO principles of Open Science
- POSI principles
- The I4OC principles that citation data should be Structured, Separable, and Open
- The FAIR data principles, that data should be Findable, Accessible, Interoperable, and Reusable

# A wide community, same values

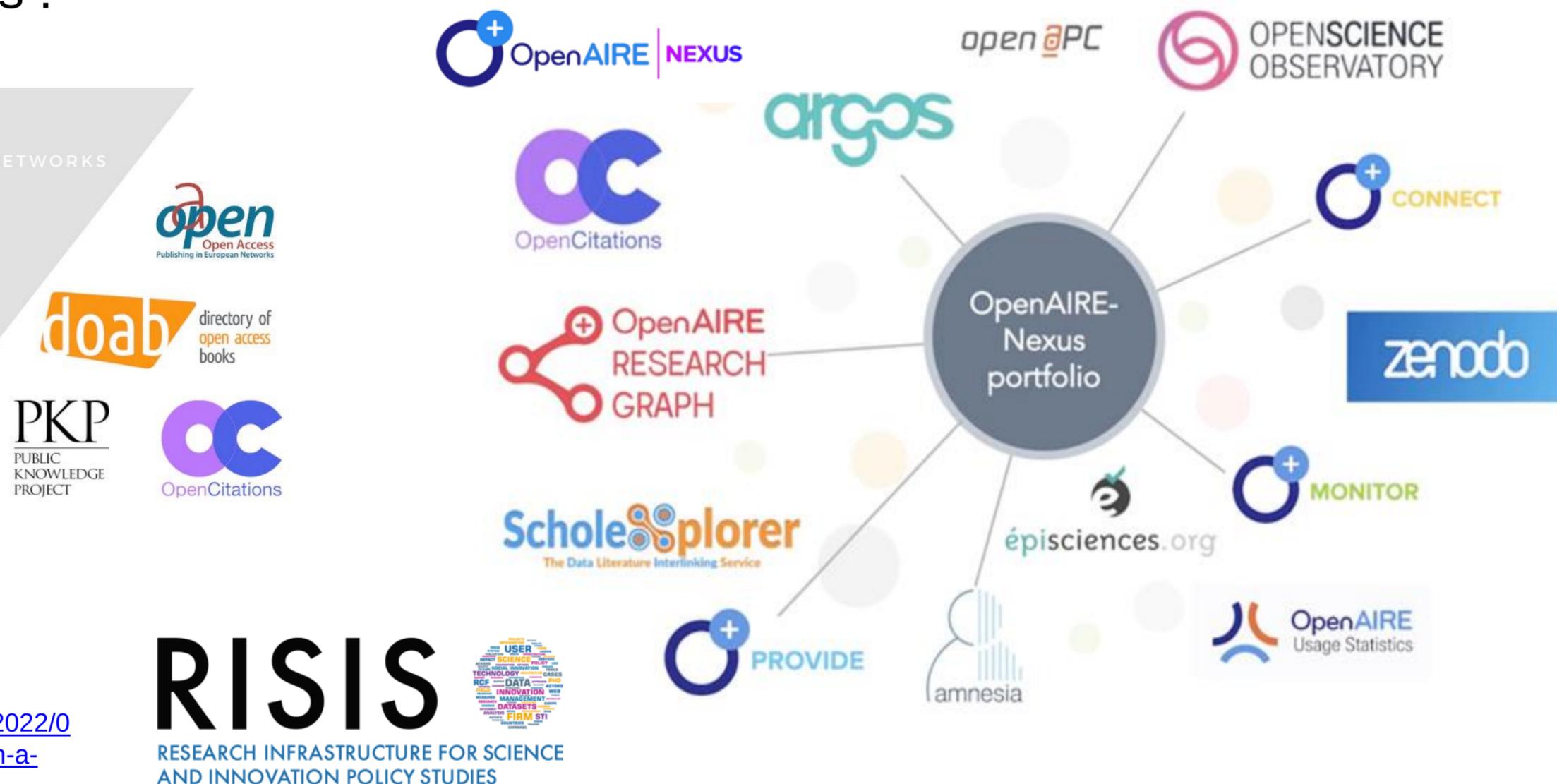
OpenCitations collaborates with projects that are part of the Open Science ecosystem and that share similar values :

DIRECTORY OF OPEN ACCESS BOOKS  
OPEN ACCESS PUBLISHING FOR EUROPEAN NETWORKS  
PUBLIC KNOWLEDGE PROJECT  
OPENCITATIONS

**SCOSS LAUNCHES SECOND FUNDING CYCLE**

Read about these essential services, their funding goals and how your institution can help support them at [www.scoss.org](http://www.scoss.org).

**SCOSS**



On our blog:

- <https://opencitations.wordpress.com/2022/01/13/five-reasons-why-2021-has-been-a-great-year-for-opencitations/>
- <https://opencitations.wordpress.com/2022/03/08/opencitations-and-ec-funding-openaire-nexus-and-risis2/>

# OpenCitations for Open Science

At the end of 2019, OpenCitations was selected by the Global Sustainability Coalition for Open Science Services (SCOSS) for their second round of crowd-funding support

SCOSS stated that OpenCitations *“aligns well with open science goals, is an innovative service, and if successful could be a game changer by challenging established proprietary citation services”*.

During FORCE 2021 conference, OpenCitations received the [Open Publishing Award](#) in the Open Data category

“It is clear that OpenCitations is one of the fundamental projects in open publishing, specifically in open scientific publishing”



# Reusability in OpenCitations

By definition of the **UNESCO Recommendations of Open Science**, science practice is “open” if it makes

*“scientific knowledge openly available, accessible and reusable for everyone”*



Re-usability is one of the main foci of OpenCitations, whose citation data satisfies all the Reuse guidelines provided by FAIR in terms of

- richness
- provenance
- usage licenses
- domain-relevant community standards.

# Data richness

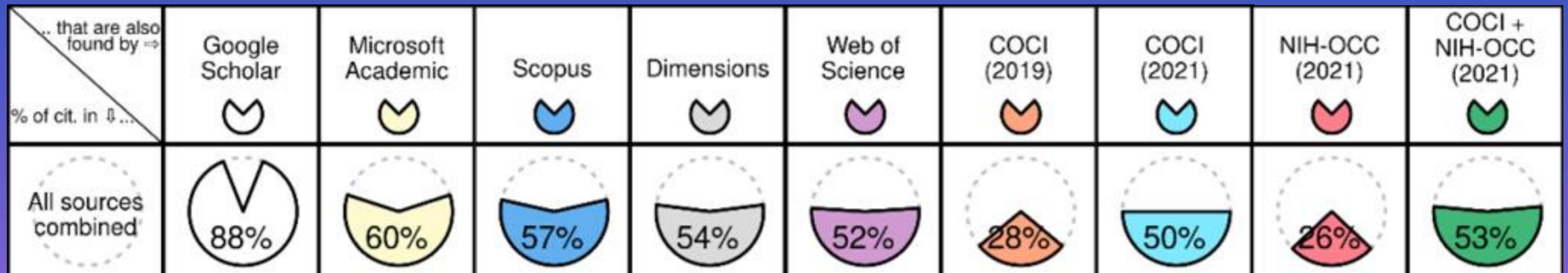
OpenCitations provides information on scholarly citations between publications covering all academic subject domains.

Our main database COCI (the OpenCitations Index of Crossref open DOI-to-DOI citations; <http://opencitations.net/index/coci/>) has recently crossed the significant threshold of one billion citations, and the last release (March 2022) contains

**1,294,283,603 citation links**

derived from open references within Crossref.

A recent independent comparison has shown that the coverage of COCI is approaching parity with those of the expensive subscription citations indexes Web of Science and Scopus



*Percentage of citations found by each database, relative to all citations (first row), and relative to the number of citations found by the other databases (subsequent rows).*

<https://opencitations.wordpress.com/2021/10/27/coverage-of-open-citation-data-approaches-parity-with-web-of-science-and-scopus/>

# Provenance information

Every item of data within OpenCitations is accompanied by **machine-readable metadata recording the provenance of the information**, including its source, the nature of any subsequent curatorial modification or correction, the dates of such actions, and the identities of the agents involved in obtaining or changing it (whether human or computational).

# Usage licenses

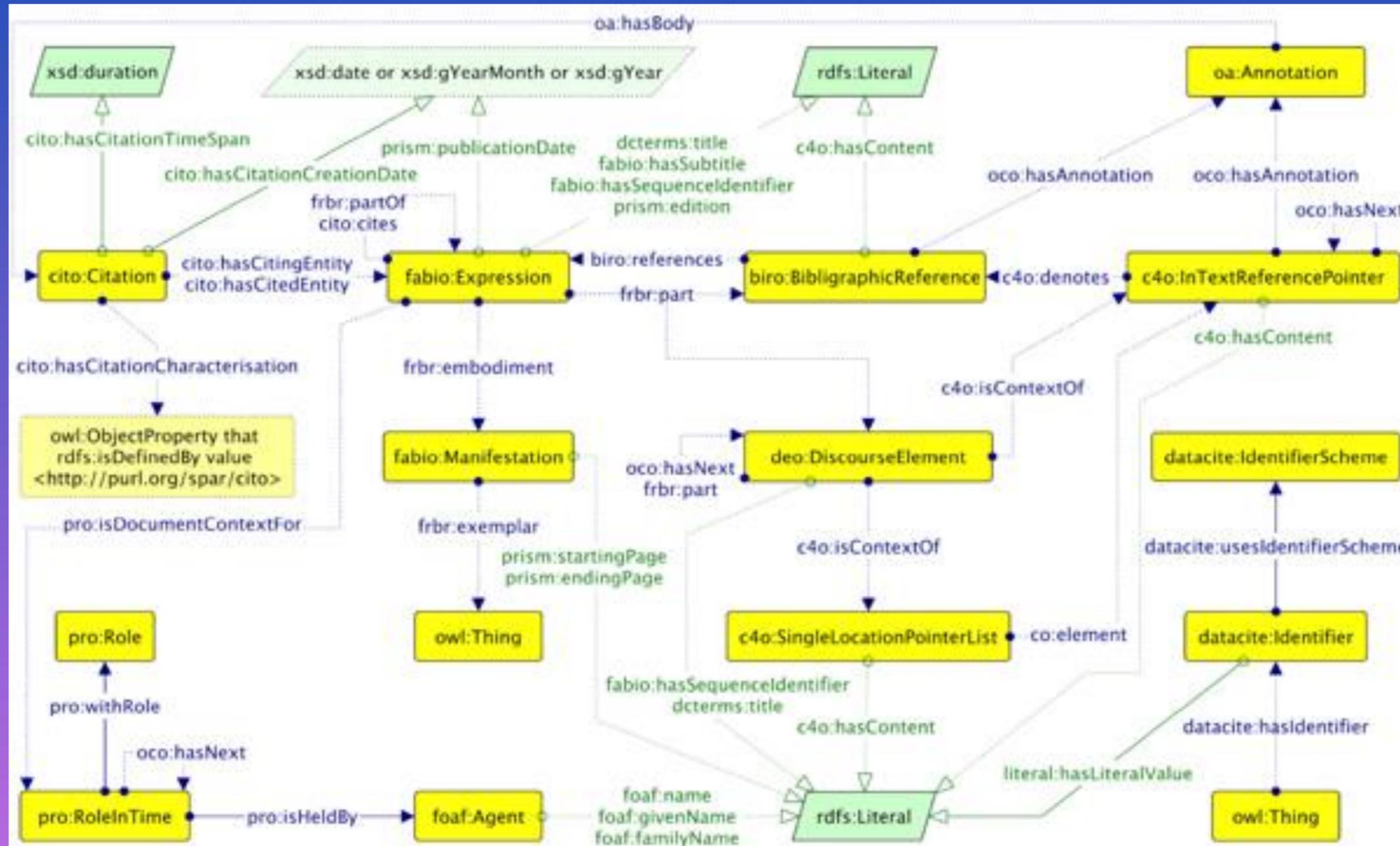
To ensure the greatest possible reusability, **all OpenCitations data are published under a Creative Commons CC0 Public Domain Waiver** that permits downloading and re-use of any nature, including commercial. Thus existing non-for-profit and commercial providers of citation indexes may themselves use OpenCitations data to expand their coverage and to enhance the value of their charged-for added-value services build over their data holdings.

# Domain-relevant community standards

## OpenCitations

- adopts Semantic Web technologies for data description, particularly the Resource Description Framework (RDF) and the Web Ontology Language (OWL);
- treats citations as first class data entities, each with a unique persistent identifier, an [Open Citation Identifier \(OCI\)](#), rather than simply as relationships between citing and cited publications;
- is thus able to assign descriptive metadata to each citation individually, for example its citation timespan, and whether or not it is a self-citation (author self-citation, journal self-citation, institutional self-citation, etc.), encoded using [CiTO, the citation typing ontology](#).

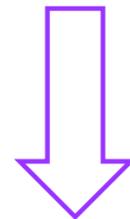
# The OpenCitations Data Model is used to structure all OpenCitations RDF data, facilitating the semantic interoperability with related data from other providers of open scholarly information



# Open source software

OpenCitations open source software, used to run all our databases and services,

- permits third parties to re-use our developments,
- ensures that the OpenCitations system can be re-created elsewhere in the unlikely event that OpenCitations itself ceases to operate



thus ensuring the **preservation and longevity** of our work and data.

OpenCitations mission is

*“to make global scholarly citation data available at zero cost and without restriction for third party analysis and reuse, in both human- and machine-readable formats”*

Would you like to embrace our mission? You can support OpenCitations activities and goals either via donation or membership  
(more info at <http://opencitations.net/membership>)

By supporting OpenCitations, you will contribute to the sustainability of a e-infrastructure for which the reusability of data is an integral part of its own mission.

Support OpenCitations

Support Open Science



**Thank you for your attention!**