

OpenCitations

OpenCitations: crossing one billion citations and beyond

Kyoto University Symposium
“Scholarly Publications as Open Data and University Libraries”

Silvio Peroni

Research Centre for Open Scholarly Metadata & Digital Humanities Advanced Research Centre (DHARC),
Department of Classical Philology and Italian Studies, University of Bologna, Bologna, Italy

Director, OpenCitations – @opencitations – <http://opencitations.net>

silvio.peroni@opencitations.net – @essepuntato – <https://essepuntato.it>



What is an open citation?

A **bibliographic citation** is a conceptual directional link from a citing entity to a cited entity, for the purpose of acknowledging or ascribing credit for the contribution made by the author(s) of the cited entity - which is defined by including a bibliographic reference in the citing article reference list and by denoting it using an in-text reference pointer (e.g. “[3]” or “(Doe et al., 2021)”).

The **citation data** related to a particular citation must include:

- the *representation* of such a conceptual directional link
- the *basic metadata* of the citing entity and the cited entity, i.e. sufficient information to create or retrieve textual bibliographic references for each of the entities

A bibliographic citation is an **open citation** when the data needed to define the citation are compliant with the following principles

Structured → Citation data expressed in one or more machine-readable formats

Separate → Citation data available without accessing the source bibliographic entity (e.g. the article or book) in which the citation is defined

Open → Citation data must be freely accessible and reusable without any restrictions

Identifiable and Available → citing and cited entities identified by using a specific persistent identifier scheme (e.g. a DOI) or a URL, and it must be possible, by resolving the identifiers of the citing and cited entities, to obtain the basic metadata of both the entities

An historical overview

The concept of open citations is strongly tied with that of **the Web**

Since **1989**, the web has drastically changed the way we think about academic publishing and science in general – they started to adopt **Web Standards** to create and deliver their products quickly and to a broader audience

Standards (e.g. JATS), guidelines (e.g. FAIR), services (e.g. REST APIs) based on Web technologies have been to improve

- the discoverability of academic products
- the research practices
- the reusability of scholarly data

Open citations are no exception

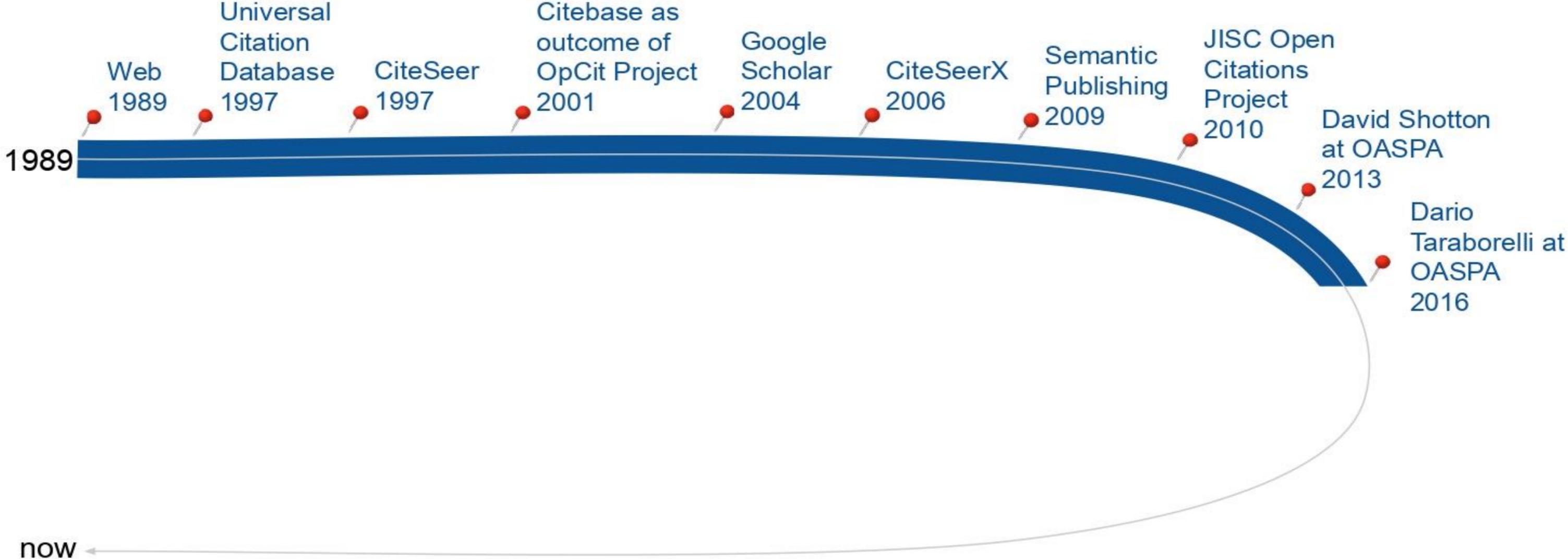
An historical overview

The first embryonal introduction of open citations is in Rober Cameron's visionary article published in 1997, "Universal Citation Database as a Catalyst for Reform in Scholarly Communication", in which he speculated about the existence of a **Universal Citation Database**

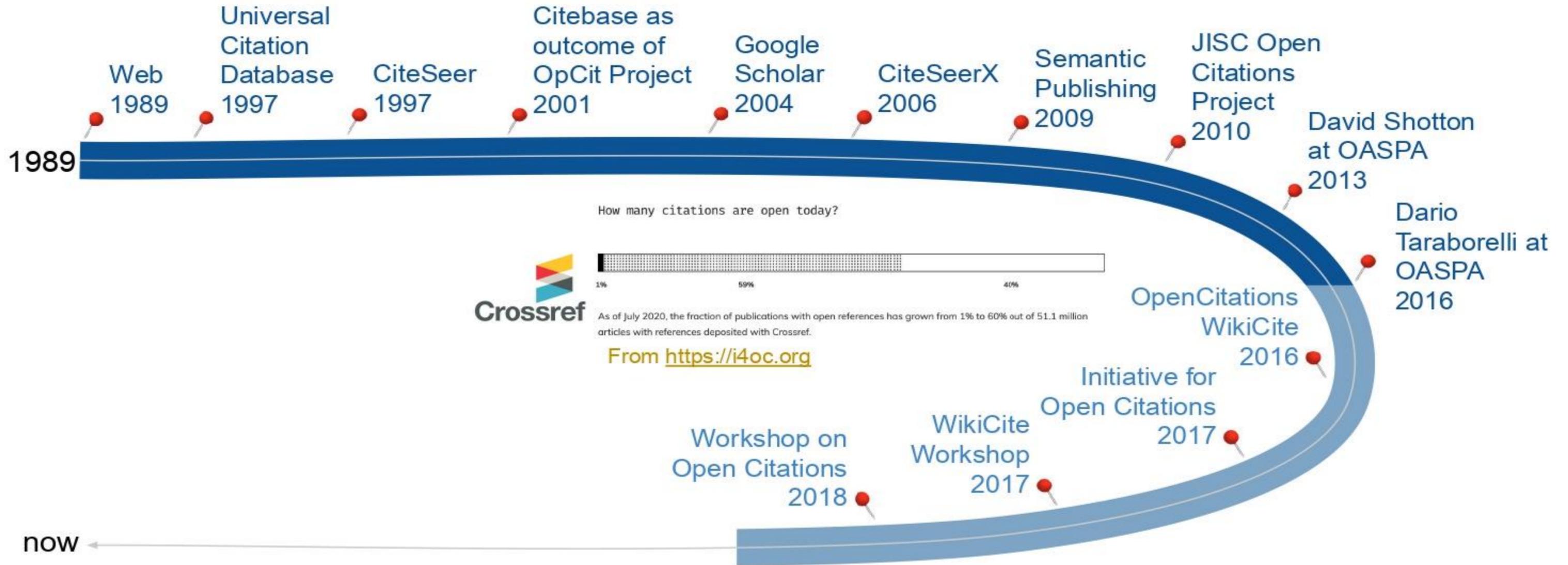
- linking every scholarly work ever written
- decentralized
- freely available over the internet
- daily updated
- all types of publications are equally visible and accessible

Cameron, R. D. (1997). A Universal Citation Database. First Monday, 2(4).
<https://doi.org/10.5210/fm.v2i4.522>

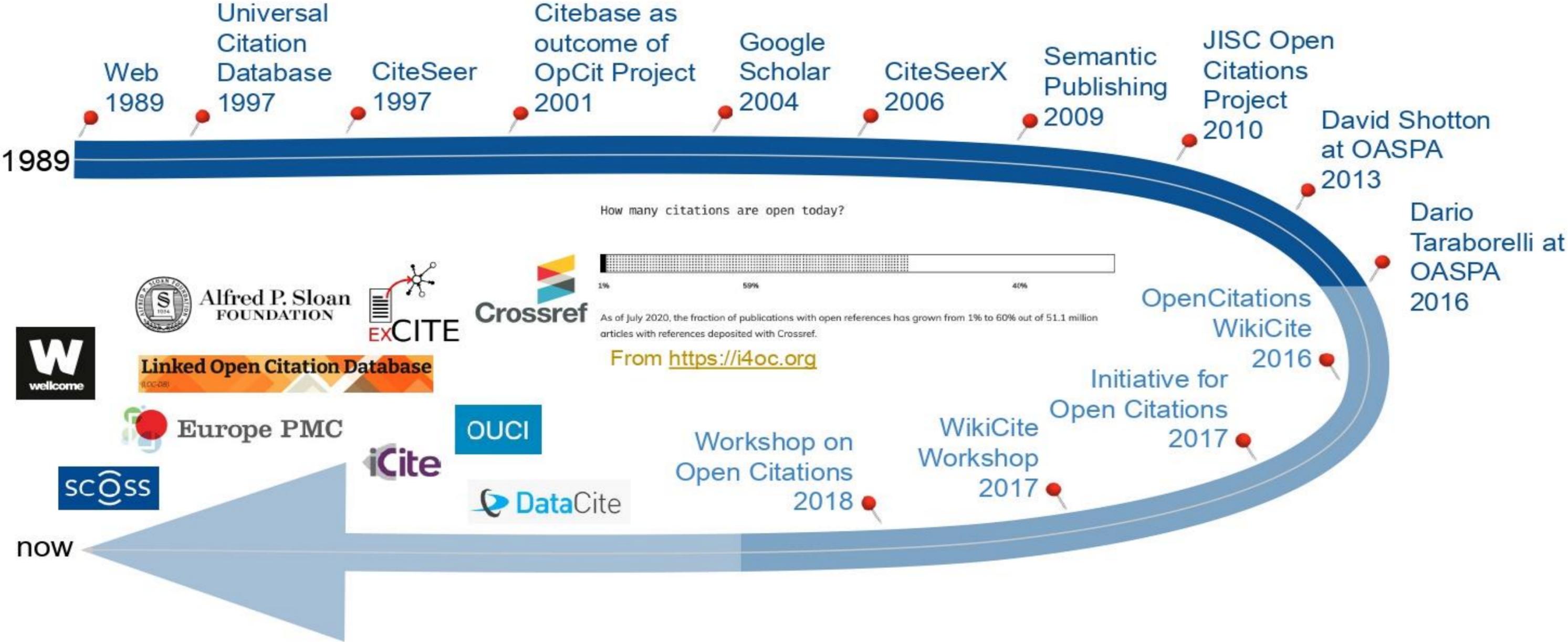
First take



Second take



Until now



OpenCitations today

[OpenCitations](#) is a scholarly infrastructure organization:

- dedicated the publication of open bibliographic and citation data, providing open metadata with a scope, depth, accuracy and provenance surpassing commercial sources
- advocacy for open scholarly metadata, particularly via **I4OC** and **I4OA**

We provide:

- data models: the [OpenCitations Data Model](#) (based on SPAR Ontologies)
- bibliographic and citation datasets: the [OpenCitations Corpus](#) and [COCI](#)
- software: [GitHub repository](#) released with open source licenses
- online services: [dumps](#), [REST APIs](#), [SPARQL endpoints](#), and [other Web interfaces](#)

Our primary database, [COCI](#) (the OpenCitations Index of Crossref open DOI-to-DOI citations) currently hosts

1,235,170,583 citation links

derived from open references within Crossref (it includes citations from **Elsevier** and the **American Chemical Society** publications)

The citations are hosted as Linked Open Data and published under the CC0 public domain waiver. These citations can be accessed via a flexible API using DOIs (Digital Object Identifiers) and OCIs (Open Citation Identifiers)

Heibi, I., Peroni, S. & Shotton, D. Software review: COCI, the OpenCitations Index of Crossref open DOI-to-DOI citations. *Scientometrics* **121**, 1213–1228 (2019).
<https://doi.org/10.1007/s11192-019-03217-6>

OpenCitations for Open Science

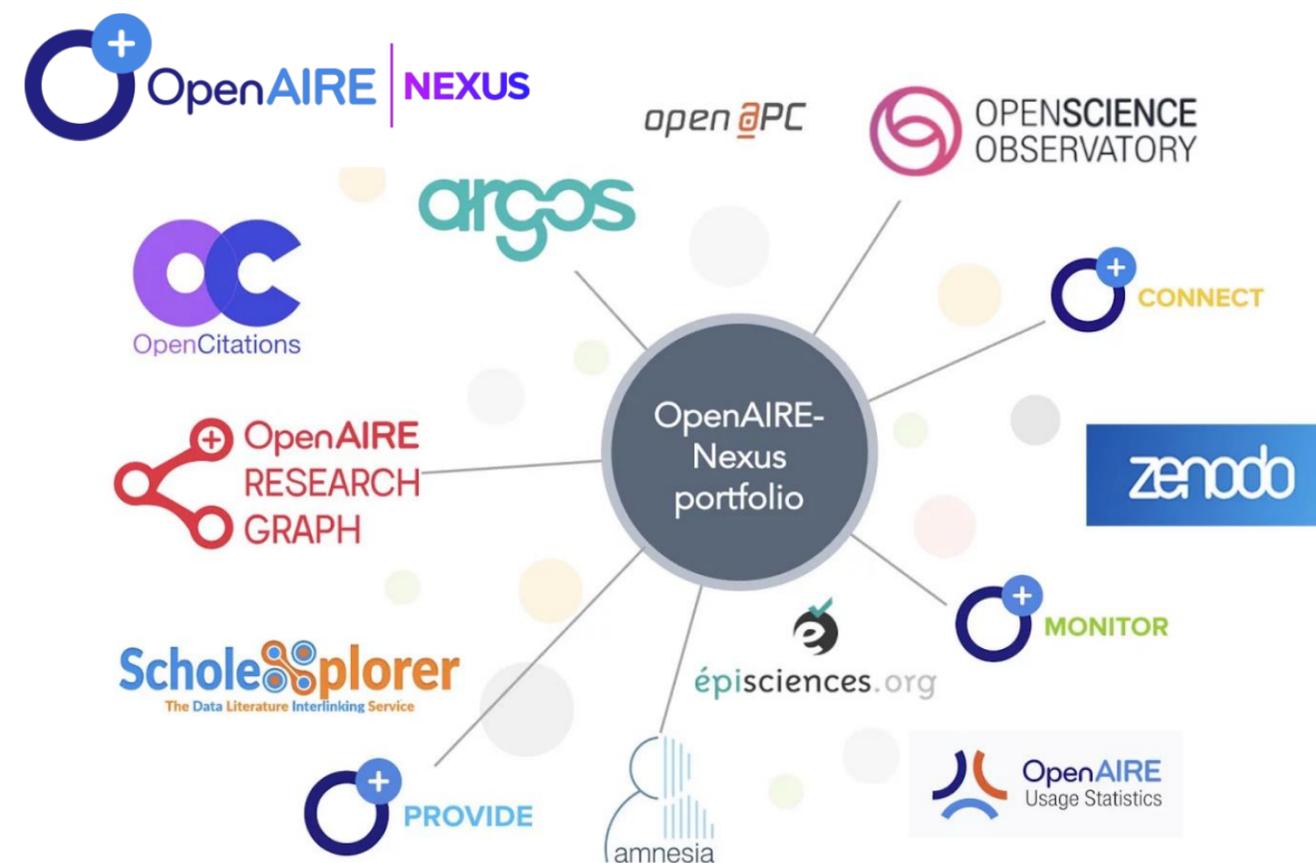
SCOSS stated that OpenCitations

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

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.





Open Citations

With so few semantic web and linked open data initiatives available in the scholarly communication landscape we wanted to present this award to a project that represents the open principles so well, so effectively and so adamantly.

<https://openpublishingawards.org/results/2021/index.html>

For OpenCitations, “open” is the crucial value and the final purpose. OpenCitations fully espouses the UNESCO founding principles of Open Science. It complies with the FAIR data principles proposed by Force11 that data should be Findable, Accessible, Interoperable and Re-usable, and with the recommendations of I4OC that citation data in particular should be Structured, Separable, and Open. By connecting publications documenting our individual scientific and cultural achievements, citation data provide the primary record of the growth of scholarly knowledge, and an explanation for how we know facts. OpenCitations consider the free availability of bibliographic citation data a necessary condition for the establishment of an open knowledge graph. Having citations “open” helps achieve a more transparent, accessible and comprehensive research practice. OpenCitations believes in the power of the scholarly community to change previous practice and, by reclaiming ownership of its own data, to create an open and inclusive future for science and research.

I would like to highlight a significant comment given on behalf of the jury panel. “At the time of writing this review, the largest database provided by OpenCitations contains more than 1.23 billion citations. Compiling this database in a license-friendly way is a feat on its own, but combine that with OpenCitations’ persistence (established 11 years ago), their active and consistent involvement with the community, and the number of works that were made possible by their effort (Google Scholar lists 1440 results), it is clear that OpenCitations is one of the fundamental projects in open publishing, specifically in open scientific publishing”.

OpenCitations espouses the following:

- UNESCO [principles of Open Science](#)
- [POSI principles](#)
- The [FAIR data principles](#)
- The [I4OC principles](#)

In compliance with these values, one of OpenCitations' main priorities is **to keep its services, software and data always without charge under open licenses (CC0 for data ISC for software) for fostering maximum reuse.**

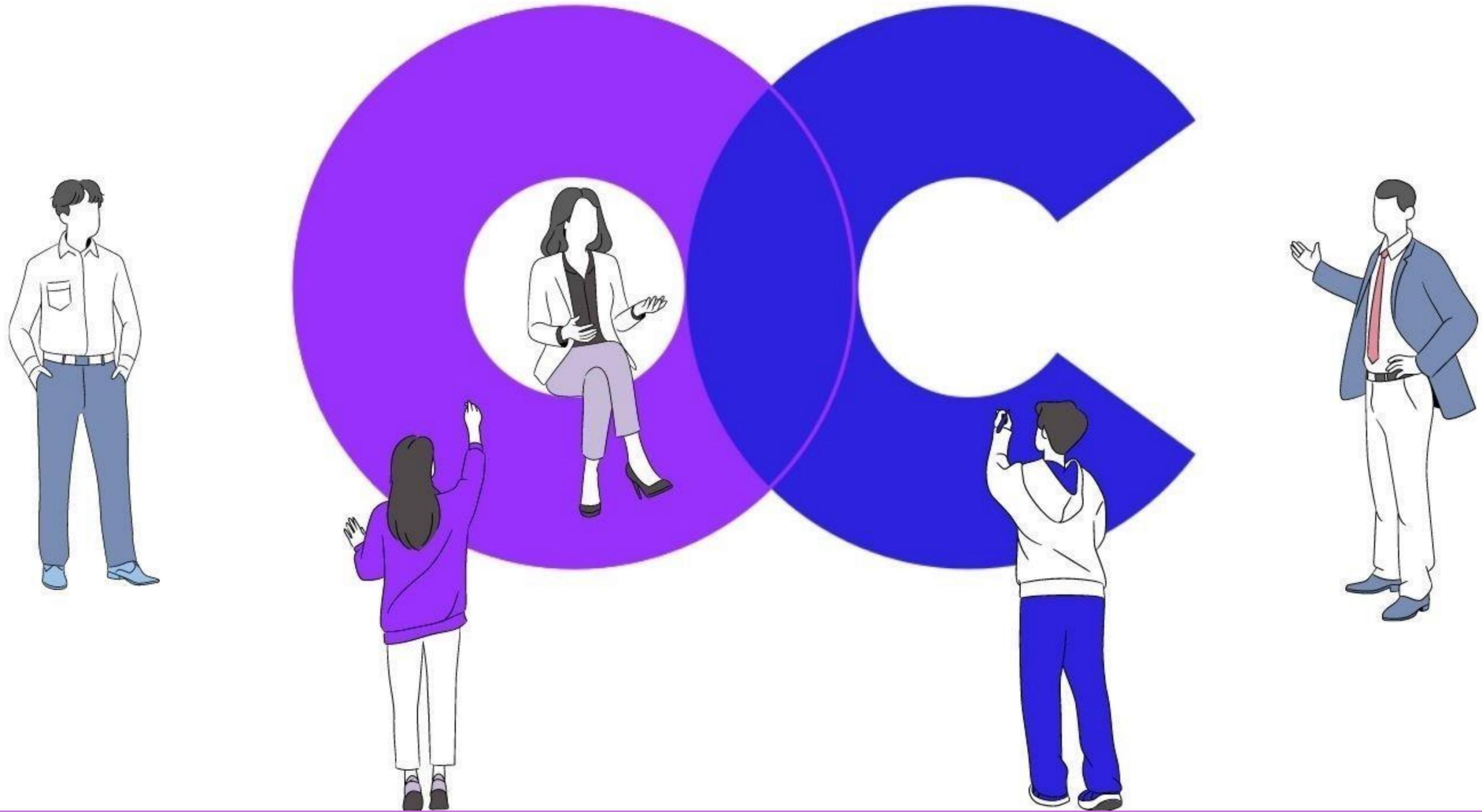
How to help

You can apply for [membership](#) and thus support us financially

- Supporting membership (€500 - €8,000 per year)
- Development membership (€8,000 - €30,000 per year)
- Strategic membership (more than €30,000 per year)

You can ask your institution to provide us your citation data and bibliographic metadata to make them available on our OpenCitations Indexes and in our new bibliographic metadata collection: OpenCitations Meta – introduced in [Peroni & Shotton \(2020\)](#). OpenCitations, an infrastructure organization for open scholarship. https://doi.org/10.1162/qss_a_00023

Together, we can make a difference and create an open and inclusive future for science and research





Thank you for your attention