



2022 SCOSS ANNUAL REPORT FORM – VERSION FOR PUBLIC SHARING

NAME OF THE REVIEWED INFRASTRUCTURE: **OpenCitations**

APPLICATION FORM: [OpenCitations Application Form](#)

YEAR IN THE CYCLE:

1/3

2/3

3/3

1. Discuss the status of the most important activities and milestones planned as specified in the Application Form (see section 6 of the Application Form). If there are important omissions or milestones that have not been achieved, please explain in the Additional comments section below.

You can keep the structure of your application or organize the milestones in main fields, such as Governance and Sustainability”, “Community Development”, “Technical Development”, “Pledges and Grants”.

Activity/Milestone 1: Collecting open references from scholarly publishers

Status:

During 2022, OpenCitations’ largest index [COCI \(the OpenCitations Index of Crossref open DOI-to-DOI citations\)](#) was able for the first time to collect the citation links which had recently been opened by [IEEE](#). In 2022, we created 5 releases of COCI. The last release of COCI (October 2022), based on open references to works with DOIs within the Crossref dump dated October 2022, contains information on 1,392,036,835 citation links from 76,072,926 bibliographic resources.

Activity/Milestone 2: Expand the coverage of open citations across the entire scholarly domain

Status:

- [OpenCitations Meta](#) database stores and delivers bibliographic metadata for all publications involved in the OpenCitations Indexes. For each publication, the metadata exposed by OpenCitations Meta includes the publication’s title, type, venue (e.g. journal name), volume number, issue number, page numbers, publication date, and identifiers such as Digital Object Identifiers (DOIs) and PubMed Identifiers (PMIDs). In addition, OpenCitations Meta includes details of the main actors involved in the document’s publication, i.e., the names of the authors, editors, and publishers, each with its own additional metadata and identifier (e.g. ORCID). Its first release (December 2022) contains information on 87,321,593 bibliographic entities, 277,750,235 authors and 2,359,301 editors (counted by their roles, without

disambiguating individuals), 710,226 publication venues, and 17,268 publishers.

- [DOCI](#), the OpenCitations Index of DataCite open DOI-to-DOI citations, is an RDF dataset containing details of all the citations that are specified in the last dump of DataCite. The citations available in DOCI are treated as first-class data entities, with accompanying properties including the citations timespan, modelled according to the OpenCitations Data Model. Its first release (December 2022) contains information on 169,822,752 citations from 358,277 bibliographic resources.
- [POCI](#), the OpenCitations Index of PubMed open PMID-to-PMID citations, is an RDF dataset containing details of all the citations from publications bearing PubMed Identifiers (PMIDs) to other PMID-identified publications, harvested from the National Institutes of Health Open Citations Collection (NIH-OCC). The citations available in POCI are treated as first-class data entities, with accompanying properties including the citations timespan, modelled according to the OpenCitations Data Model. Its first release contains 717,654,703 citations from 26,024,862 bibliographic resources.

Activity/Milestone 3: New developers

Status:

During 2022, OpenCitations development team welcomed two new members, for a total of four developers now working daily on OpenCitations:

- Arianna Moretti (University of Bologna) was appointed as a new software developer to work on the improvement and release of OpenCitations new indexes POCI and DOCI.
- Ivan Heibi (University of Bologna) was appointed to be responsible for the OpenCitations technical infrastructure, with the role of technical supervision and maintenance.

Our other two developers are Arcangelo Massari (software developer, hired in December 2021); Giuseppe Grieco (software and systems developer, hired in June 2021).

In December, OpenCitations published a [job advert](#) for a Research Fellow position dedicated to the management and improvement of the OpenCitations technical infrastructure, to be appointed by the Research Centre for Open Scholarly Metadata at the University of Bologna from April 2023.

Activity/Milestone 4: Governance and legal administration

Status:

During 2022, we worked to further develop the open governance structure of OpenCitations. For administrative convenience, OpenCitations is managed by the Research Centre for Open Scholarly Metadata at the University of Bologna. Currently, the organizational bodies included in the OpenCitations governance are: the Directors, the International Advisory Board, and the Council, regulated and described in the documents "[Organizational Bodies and Finances](#)", and "[Rules for Membership](#)". In November 2022, in anticipation of the end of SCOSS sponsorship, we slightly edited



the criteria of membership to OpenCitations, by making it possible for institutions to apply for one-year-long membership.

During 2022, OpenCitations held four International Advisory Board Meetings and an annual Council Meeting.

The OpenCitations “Governance Evolution” working group started meeting in May 2022, involving OpenCitations’ director Silvio Peroni, the OpenCitations Research Manager Claudio Fabbri, and some members of the International Advisory Board. The group met four times during 2022, among other things working to define more clearly the role of the University of Bologna in OpenCitations’ governance and internal structure. The working group’s strategic role will be improved during 2023.

Activity/Milestone 5: Community Development

Status:

Being a community-based infrastructure organization, OpenCitations dedicates part of its activities to community outreach. In order to identify and reach out to wider communities and new stakeholders, in July 2022 OpenCitations started a Community Building Working Group, involving OpenCitations director David Shotton, the Communications Director Chiara Di Giambattista, and members from the International Advisory Board. In 2022 the working group met three times, and will continue its strategic work in 2023.

In August, we published three documents to publicly define and present OpenCitations’ mission, unique strengths and next developmental steps:

- [OpenCitations Mission Statement](#)
- [The Uniqueness of OpenCitations](#)
- [OpenCitations – Present Status and Future Plans.](#)

We enhanced the communication of OpenCitations’ news using OpenCitations Blog, social media and communications partnerships.

Social Media communication: We increased OpenCitations interactions on Twitter and LinkedIn profile. Social Media platforms were used to share and promote the blog contents and for brief communications.

We created an OpenCitations’s profile on a new social media platform, [Mastodon](#).

OpenCitations blog: in August 2022, we moved OpenCitations blog to the academic blogs platform Hypotheses.org. The new platform made it possible to renovate and improve our blog, with a new appearance and better tools. OpenCitations’ renovated blog is at <https://opencitations.hypotheses.org/>



OpenCitations blog hosted guest posts, and promoted some post series:

- it hosted two technical posts by OpenCitations developers Ivan Heibi and Arcangelo Massari, including one tutorial as part of the post series #techexplained:
<https://opencitations.hypotheses.org/2940>
- it reposted the blog post “Transparency meets open citations” from OpenCitations Director Silvio Peroni’s blog *QUERTY: musings from the rabbit hole*.
<https://opencitations.hypotheses.org/2890>
- it promoted the releases of the new indexes POCI and DOCI, and OpenCitations Meta with the post series #OC2022releases”.

The five most recent posts are as follows:

- [Discover POCI, the index of open citations from PubMed](#)
- [A technical overview of OpenCitations Meta](#)
- [Discover DOCI, the index of open citations from DataCite](#)
- [Please help with OpenCitations’ entry in Wikipedia](#)
- [Follow OpenCitations on Mastodon](#)

Communication partnerships:

As part of the OpenAIRE Nexus project, the OpenAIRE communication team supported OpenCitations in many outreach activities, including the publication of a guide, a factsheet, a public roadmap, a blog post on OpenAIRE blog and a post on the OpenAIRE September 2022 newsletter;

SCOSS helped to disseminate OpenCitations news by retweets and by sharing the 2021 report in the May newsletter;

OpenCitations dedicated three blog posts to promote its partnerships and collaborations :

- [Five reasons why 2021 has been a great year for OpenCitations](#)
- [OpenCitations and EC funding: OpenAIRE Nexus and RISIS2](#)
- [Two years of achievements within the ‘SCOSS family’ \(and it’s not over yet!\)](#)

During 2022, OpenCitations participated in numerous events and conferences that helped reach out to a wider community:

EAHIL Evaluations and Metrics Webinar 2022

February 2, 2022

The session was organised by the Evaluation and Metrics special interest group from the European Association for Health Information and Libraries (EAHIL), and dedicated to the presentation of OpenCitations data sets, tools, services, and activities.

Slides available at:

Peroni, Silvio. (2022, February 2). OpenCitations: an independent infrastructure organization for open citation data. EAHIL Evaluation and Metrics SIG Webinar. Zenodo.

<https://doi.org/10.5281/zenodo.5949271>

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

February 15, 2022

A session entitled ‘OpenCitations: crossing one billion citations and beyond’ to present OpenCitations to the Kyoto University librarian community. Slides available at:

Peroni, Silvio (2022) [第一部 : オープン・サイテーション]
OpenCitations: crossing one billion citations and beyond. Kyoto University Symposium “Scholarly Publications as Open Data and University Libraries”. <http://hdl.handle.net/2433/268886>

KB Webinar Series “Future Perspectives of Open Citation Data”

March 16, 2022

A session entitled ‘Enabling transparency in research assessment’ with a presentation about OpenCitations’ role in research assessment . Slides available at:

Peroni, Silvio. (2022, March 16). Enabling transparency in research assessment. Future Perspectives of Open Citation Data, Virtual Event. Zenodo.

<https://doi.org/10.5281/zenodo.6365246>

DataCite and Make Data Count webinar “FORAGE: the hunt for existing data citations”

March 17, 2022

A session entitled ‘Open citations for enabling transparent metrics’, as part of a webinar series dedicated to surfacing data citations beyond publisher workflows. Slides available at:

Peroni, Silvio. (2022, March 17). Open citations for enabling transparent metrics. FORAGE: the hunt for existing data citations, Virtual. Zenodo.

<https://doi.org/10.5281/zenodo.6386139>

ENG4GED webinar

April 14, 2022

A webinar session entitled 'OpenCitations: an infrastructure for open bibliographic and citation data' as part of a webinar series organized by B4DS - Business Engineering for Data Science (University of Pisa). Q&A available at:

[ENG4GED - OpenCitations: an infrastructure for open bibliographic and citation data](#)

19th Extended Semantic Web Conference (Hersonissons, Greece)

June 2, 2022

Conference presentation of the paper 'A programming interface for creating data according to the SPAR Ontologies and the OpenCitations Data Model', written by Simone Persiani (University of Bologna), Marilena Daquino (University of Bologna) and Silvio Peroni (University of Bologna). Slides available at:

Peroni, Silvio. (2022, June 2). A programming interface for creating data according to the SPAR Ontologies and the OpenCitations Data Model. Zenodo.

<https://doi.org/10.5281/zenodo.6606656>

17th International Digital Curation Conference

June 14, 2022

A virtual presentation by Ivan Heibi to present the short abstract 'OpenCitations, an open e-infrastructure to foster maximum reuse of citation data'. Slides available at:

Di Giambattista, Chiara, Heibi, Ivan, Peroni, Silvio, & Shotton, David. (2022, June 7).

OpenCitations, an open e-infrastructure to foster maximum reuse of citation data. Zenodo.

<https://doi.org/10.5281/zenodo.6622077>

ULITE workshop 2022

June 23, 2022

Keynote speech 'OpenCitations. A short introduction', by Silvio peroni, given during the ULITE workshop held in conjunction with JCDL 2022. Slides available at:

Peroni, Silvio. (2022, June 24). OpenCitations – A Short Introduction. ULITE Workshop 2022 (ULITE-ws), Virtual. Zenodo. <https://doi.org/10.5281/zenodo.6719497>

Interactive demo session by Ivan Heibi and Arcangelo Massari, entitled ‘How to structure citations data and bibliographic metadata in the OpenCitations accepted format’. Proceedings available at <https://ceur-ws.org/Vol-3220/>

ResearchHub Online SciCon 2022

July 23, 2022

A virtual conference dedicated to the intersection between Metascience and Decentralized Science. The slides of OpenCitations presentation ‘Clearing the fog on metascience – or, how to enable reproducible studies through open citations’ are available at:

Peroni, Silvio. (2022, July 23). Clearing the fog on metascience - or, how to enable reproducible studies through open citations. ResearchHub SciCon 2022 (SciCon 2022), Online. Zenodo. <https://doi.org/10.5281/zenodo.6892143>

TripleEU Booksprint (Konstancin, Poland)

September 7, 2022

An on-site event, organized by the TRIPLE consortium and coordinated by IBL PAN, focusing on a limited number of key types of metadata for a scholarly publication and assess it in the context of open metadata initiatives and most pressing challenges in the Social Sciences and Humanities. The slides of Silvio Peroni’s talk ‘Transparency meets open citations: how to enable reproducible studies through open citations’ are available at:

Peroni, Silvio. (2022, September 7). Transparency meets open citations: how to enable reproducible studies through open citations. Zenodo. <https://doi.org/10.5281/zenodo.7063176>

TPDL Conference (Padua, Italy)

September 28, 2022

Ivan Heibi presented the paper ‘Enabling Portability and Reusability of Open Science Infrastructures’ by OpenCitations development team (Giuseppe Grieco, Ivan Heibi, Arianna Moretti, Arcangelo Massari, Silvio Peroni). Paper available at:

Grieco, G., Heibi, I., Massari, A., Moretti, A., Peroni, S. (2022). Enabling Portability and Reusability of Open Science Infrastructures. In: , et al. Linking Theory and Practice of Digital Libraries. TPDL 2022. Lecture Notes in Computer Science, vol 13541. Springer, Cham. https://doi.org/10.1007/978-3-031-16802-4_36

EOSC Symposium 2022 (Prague, Czech Republic)

November 14, 2022

Arianna Moretti participated in the annual EOSC Symposium (part of the calendar of events of the Czech presidency of the Council of the EU) with the poster presentation 'OpenCitations, an infrastructure organization that harvests and publishes FAIR and open metadata describing the world's academic publications and the scholarly citations'. Poster at:

Chiara Di Giambattista, Arianna Moretti, & Silvio Peroni. (2022). OpenCitations, an infrastructure organization that harvests and publishes FAIR and open metadata describing the world's academic publications and the scholarly citations. EOSC Symposium 2022, Prague. Zenodo. <https://doi.org/10.5281/zenodo.7341057>

In 2022, OpenCitations organized two events:

- On October 5, we held the 4th **Workshop on Open Citations and Open Scholarly Metadata**, a 3-hour webinar for researchers, scholarly publishers, funders, policy makers, and opening citations advocates, interested in the creation, reuse, and improvement, of open citation data and open scholarly metadata, with invited speakers. The materials from the call for contributions and invited talks are available on Zenodo: <https://zenodo.org/communities/wooc2022/?page=1&size=20> .

The video record is available on YouTube: https://www.youtube.com/watch?v=82h_YVzA_CU .

- On December 12 we held the **OpenCitations Annual Webinar** for the OpenCitations Council members, an internal event during which the OpenCitations directors provided an overview of OpenCitations and presented the latest updates and future developments.

In 2022, OpenCitations was involved in two funded applications and R&D projects and started two new collaborations:

OpenAIRE Nexus

The funded [OpenAIRE Nexus project](#) started on 1 January 2021 and will end in June 2023. OpenCitations was involved (through the University of Bologna) in the preparation of a

European consortium bid, with other ten partners, several of which were already working actively in OpenAIRE, in response to an H2020 call ([INFRAEOSC-07-2020 - Increasing the service offer of the EOSC Portal](#)). The main aim of the project is to create a framework of services for assisting in publishing research, monitoring its impact, helping promote its discovery, and integrating it into the European Open Science Cloud (EOSC) for the benefit of the open science community worldwide. OpenCitations' role in this is to provide open bibliographic citations as part of the open data components of OpenAIRE and the EOSC. The support received from OpenAIRE Nexus made it possible to appoint Arcangelo Massari, who worked on the development and release of OpenCitations Meta.

OpenCitations is now included in the OpenAIRE Service Catalogue and in the EOSC Marketplace.

RISIS2 Project

[The Research Infrastructure for Science and Innovation Policy Studies \(RISIS\)](#) is a project funded by the European Union under [Horizon2020](#) Research and Innovation Programme Grant Agreement. RISIS2 gathers 18 partners aiming to transform the field of STI studies into an advanced research community. OpenCitations was involved in the project in December 2021 through a third-party agreement. The project ended in December 2022, with the release of OpenCitations Meta and the new index DOCI as outcomes.

JISC

[JISC](#) is a United Kingdom not-for-profit company that provides network and IT services and digital resources in support of further and higher education institutions and research as well as not-for-profits and the public sector. In November 2022, OpenCitations and Jisc concluded an agreement to include OpenCitations in the [Jisc Collections](#), thereby permitting UK institutions to support OpenCitations through JISC.

GraspOS

At the end of 2022, OpenCitations signed an agreement with the [GraspOS](#) project (which started on January 1, 2023), to be integrated into an open and trusted federated infrastructure for next-generation research metrics and indicators. GraspOS aims to build and operate a data infrastructure to support the policy reforms and pave the way towards a responsible research assessment system that embeds OS practices and accelerates its adoption in Europe. GraspOS will focus on extending the EOSC ecosystem with tools and services that will facilitate monitoring the use and uptake of various types of research services and outputs (publications, datasets, software), and will catalyse the implementation of policy-level rewards to foster OS practices.