
OpenAIRE services for Open Science

Alessia Bardi^{*1}, Miriam Baglioni^{*1}, Argiro Kokogiannaki², Pedro Principe³, André Vieira³, Paolo Manghi⁴, and Erzsébet Tóth-Czifra⁵

¹Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo" – Italy

²ATHENA - Research and Innovation Center in Information, Communication and Knowledge Technologies – Greece

³University of Minho [Braga] – Portugal

⁴National Research Centre – Italy

⁵DARIAH – Germany

Abstract

According to Open Science (OS) principles, all scientific products generated by research activities (e.g. scientific literature, research data, software, experiments) should be made available as soon as possible “under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods”. [1]

The effective implementation of Open Science is today hindered by several cultural and technical barriers. Publishing and reward policies are still focused only on research literature (books, journal articles, etc.). Attention to research data is growing in some communities of practice, while the publishing of research packages representing digital experiments is still not a usual procedure. The digital tools and services provided by Research Infrastructures and used by researchers are not connected to the scholarly communication ecosystem: publishing is therefore still a manual action, typically performed by researchers at the end of a research activity. Research products are made available on several scholarly communication data sources (e.g. repositories, archives, which may be institutional repositories, community-specific archives, generic repositories), making it hard to keep up-to-date links between products (e.g. research data re-used by several studies) and to identify those that are relevant for a specific research community or that have been produced using services and tools provided by a given research infrastructures.

The OpenAIRE initiative (www.openaire.eu) is a legal entity representing a partnership of more than 50 institutions working to promote and support a sustainable implementation of Open Access and Open Science policies for reproducible science, transparent assessment and omni-comprehensive evaluation.

It supports the implementation and alignment of Open Science policies at the international level by developing and promoting the adoption of global open standards and interoperability guidelines [2] to realize a sustainable, participatory, trusted, scholarly communication ecosystem, open to all relevant stakeholders (e.g. research communities, funders, project coordinators) and capable of engaging society and foster innovation.

Thanks to the network of National Open Access Desks (NOADs), OpenAIRE supports the

*Speaker

implementation of Open Science at the local and national level, supporting researchers, project coordinators, funders and policy makers with training and support activities like workshops and webinars.[3]

Furthermore, OpenAIRE features a technical infrastructure that materializes an open, de-duplicated, participatory metadata research graph of interlinked scientific products (including research literature, datasets, software, and other types of research products like workflows, protocols and methods), with access rights information, linked to funding information, research communities and infrastructures. The graph is materialized by collecting more than 100 millions of metadata records from more than 9,000 scholarly data sources world-wide.[4] In addition to the information collected from scholarly data sources, the graph includes metadata information and links between products that are *(i)*asserted (i.e. claimed) by users of the OpenAIRE Explore portal via the "Linking" functionality, and *(ii)*inferred by full-text and metadata mining algorithms. In particular, OpenAIRE algorithms enrich the graph with:

- Links from research products (literature, data, software or products of other types) to project fundings;
- Links from research products to research infrastructures;
- Links between research products (e.g. from literature to research software and data);
- Subjects and keywords based on standard classification schemes;
- Authors' affiliations;
- Countries for which the product is relevant;
- Research communities for which the product is relevant;
- Citations.

The resulting graph is called the *OpenAIRE Research Graph* and it is openly available under CC-BY license[5] to programmatic clients via the Develop API: <https://develop.openaire.eu>. Added-value services are also built on top of the graph to offer Open Science services to different stakeholders of the scholarly communication ecosystem.

Researchers can deposit any kind of research product on Zenodo (www.zenodo.org), the OpenAIRE "catch-all" repository, exploit the OpenAIRE Explore portal (<https://explore.openaire.eu>) to discover research products and check if there is an Open Access version available. Moreover, the Explore portal offers a reporting tool to *project coordinators*, who can also view statistics about the research products funded by a given project.

Content providers, like institutional repositories, have access to a plethora of services that support them in the implementation of interoperability guidelines, in the collection of usage statistics and for metadata enrichment via the Content Provider Dashboard (<https://provide.openaire.eu>).

Research communities can benefit from a dedicated gateway where researchers can find all the products that are relevant to a given community and have easy access to Open Science publishing tools thanks to the Research Community Dashboard (<https://connect.openaire.eu>).

Funders and research infrastructures can monitor their research impact, the compliance to their Open Access mandates, and the uptake of Open Science publishing practices among the researchers they funded via the OpenAIRE Monitor portal (<https://monitor.openaire.eu>).

Acknowledgments: This work has been co-funded by the OpenAIRE-Advance project (Horizon 2020) under Grant number 777541.

FOSTER definition of Open Science: <https://www.fosteropenscience.eu/foster-taxonomy/open-science-definition>

OpenAIRE Guidelines for Content Providers: <https://guidelines.openaire.eu>

OpenAIRE support and training on Open Science: <https://www.openaire.eu/support>

In the OpenAIRE beta infrastructure -<https://beta.explore.openaire.eu>. The production infrastructure (<https://explore.openaire.eu>) collects more than 30 millions of metadata records from more than 13,000 scholarly data sources.

Some data sources aggregated by OpenAIRE do not allow to re-distribute metadata in CC0; the availability of another version of the graph including only information that can be redistributed in CC0 is in OpenAIRE's plans.

Keywords: Open Access, Open Science, scholarly communication, research impact