

# OpenAIRE infrastructure and services: advancing Open Science

*Pedro Príncipe, University of Minho, pedroprincipe@sdum.uminho.pt*

## Session Type

- Presentation

## Abstract

OpenAIRE has established itself as a key and sustainable infrastructure for giving access to Open Access publications in Europe and beyond, progressively providing access to datasets, software and other research artefacts. From its outset, OpenAIRE has pursued a service-driven design to engage all stakeholders and the current service portfolio (covering all e-Infrastructure layers) targets a variety of users, namely researchers, content providers, funders and research communities.

OpenAIRE infrastructure is currently able to deliver a set of relevant services for content providers managers. The OpenAIRE Literature Broker Service is a tool operating on top of the OpenAIRE information graph and supports repository managers with a web dashboard where they can monitor all their repositories and can view the enrichments suggested by the information graph.

Funders can currently benefit from a set of services to monitor research outputs and impact and to integrate a body of resources in their ecosystems. OpenAIRE has now successfully applied the model and services developed for the European Commission to other funders, mainly from European Union.

OpenAIRE is working closely with existing Research Infrastructures and research communities to extend its service portfolio by introducing two new services implementing the concept of “Open Science as a Service”: Research Community Dashboard and Catch-All Broker Service.

OpenAIRE-Advance, the new phase of OpenAIRE infrastructure, continues the mission of OpenAIRE to support the Open Access and Open Data mandates in Europe. By sustaining the current infrastructure, comprised of a human network and technical services, it consolidates its achievements while working to shift the momentum among its communities to Open Science, aiming to be a trusted e-Infrastructure within the realms of the European Open Science Cloud.

## Conference Themes

- Content - research data, digital preservation, persistent urls, archiving
- Projects - sustainability of projects beyond the grant, maturing communities
- Infrastructure/Integrations - integrations between systems, changing technical environments
- Policy - national, international, local and community policy and decisions

## Keywords

OpenAIRE infrastructure, Open Access, Repositories, European Open Science Cloud, Open Science policies

## Audience

Policy makers, librarians, repository managers, research managers, librarians

## **Background**

OpenAIRE fosters the social and technical links that enable Open Science in Europe and beyond. OpenAIRE is the European initiative for an Open Access Infrastructure for Research in Europe which supports open scholarly communication and open science, and access to the research output of European funded projects.

OpenAIRE gathers content from a network of institutional and disciplinary repositories around Europe and beyond. To put this into practice, an integrated suite of guidelines has been developed so that data sources, namely literature and data repositories, open access publishers and current research information systems, can be harvested and their contents made more visible. An overarching portal gives access to the harvested open access publications, EU-funded scholarly publications and datasets, registered in over 1000 participating content providers. This massive collection enables OpenAIRE to enrich the collected dataset as a whole and provide services such as project publications and datasets lists, project monitoring tools for depositing, usage statistics and interchanging enriched data.

OpenAIRE is building a comprehensive infrastructure that covers all types of scientific output and has established an infrastructure to harvest, enrich and store the metadata of scientific data. Cross-links from publications to data and funding schemes are supported. This interlinking of research objects has implications on the optimisation of the research process, allowing the sharing, enrichment and reuse of data. It positions open access publications in the wider research context, i.e. through linking to funding information, associated datasets, software and other types of research outputs.

OpenAIRE infrastructure is implementing and aligning Open Science policies across Europe and the world and delivering a set of services to embed Open Science into different stakeholder's daily workflows, such as repository managers, research managers and researchers.

## **Content**

### **OpenAIRE infrastructure and the literature broker service**

The OpenAIRE infrastructure collects metadata records from various different data sources (journals, literature repositories, funders, data repositories) and derives from them objects and relationships that form the OpenAIRE information space graph. The OpenAIRE infrastructure performs a series of actions involving metadata and interoperability protocols that allow content enrichment in local repositories and collections. Data sources that provide content to OpenAIRE and are interested in augmenting their local collections benefit from this information graph in a number of ways. This is particularly true for institutional repositories, whose mission is to build a complete collection of the scientific publications of their affiliated authors. The advantage of the data enrichment provided by the information graph is that all articles by affiliated authors can be available in their institutional collections, and metadata are as complete and up-to-date as possible. The OpenAIRE Literature Broker Service is a tool operating on top of the OpenAIRE information graph and supports repository managers with a web dashboard where they can monitor all their repositories and can view the results suggested by the information graph. On the dashboard, repository managers can subscribe to notifications for enrichments and additions, relevant to their repository, appearing at the OpenAIRE information space graph. The OpenAIRE Literature Broker's notifications will help repository managers to learn about publication objects in OpenAIRE that do not appear in their collection but might belong to it, and be aware of additional properties or relationships relevant to publication objects in their collection.

### **OpenAIRE services for funders and the role of repositories**

A major activity in OpenAIRE is the linking of research results to funding based on the network of repositories. OpenAIRE supports the Open Access policies of the European Commission by gathering and displaying the Open Access publications of the FP7 and Horizon 2020 funding streams. After developing

these services for the EC and ERC, OpenAIRE has now successfully applied this model and its services to include additional funders, with others in development. To collect funders' data, OpenAIRE only requires a very limited set of metadata. After aggregating the data, OpenAIRE enables advanced monitoring (including compliance with Open Access policies), reporting and analysis of research impact and trends. Funders can assess the impact of their funding by viewing advanced statistics on research outputs and the funding programme/stream/project from which they derive (including co-funded research results and research trends). OpenAIRE enriches project metadata by adding new information and links to related publications, data, authors, etc, collected from the OpenAIRE compatible repositories. OpenAIRE then publishes this enriched metadata under open licenses so funders and repositories can access and re-use it via the OpenAIRE portal and a direct machine to machine application programming interface (<http://api.openaire.eu>). Providing their data to OpenAIRE, funders benefit from added value features like automatic reporting of project outputs on the EC's Participants portal.

### **OpenAIRE services in support of Open Science as-a-Service**

The effective implementation of Open Science calls for a scientific communication ecosystem capable of enabling the "Open Science publishing principles" of transparency and reproducibility. Such ecosystem should provide tools, policies, and trust needed by scientists for sharing/interlinking (for "discovery" and "transparent evaluation") and re-using (for "reproducibility") all research products produced during the scientific process, e.g. literature, research data, methods, software, workflows, protocols, etc. OpenAIRE fosters Open Science by advocating its publishing principles across Europe and research communities aiming to provide Research Infrastructures (RIs) with the services required to bridge the research life-cycle they support. OpenAIRE is fostering the establishment of reliable, trusted, and long lasting RIs by compensating the lack of OS publishing solutions and providing the support required by RIs to upgrade existing solutions to meet Open Science publishing needs (e.g. technical guidelines, best practices, OA mandates). To this aim, OpenAIRE is working closely with existing RIs to extend its service portfolio by introducing two new services implementing the concept of "Open Science as a Service" (OSaaS):

- The Research Community Dashboard. Thanks to its functionality, scientists of RIs can
  - find tools for publishing all their research products, such as literature, datasets, software, research packages, etc. (provide metadata, get DOIs, and ensure preservation of files),
  - interlink such products manually or by exploiting advanced mining techniques, and
  - integrate their services to automatically publish metadata and/or payload of objects into OpenAIRE. As a consequence, scientists populate and access an information space of interlinked objects dedicated to their RI, through which they can share any kind of products in their community, maximise re-use and reproducibility of science, and outreach the scholarly communication at large.
- The Catch-All Broker Service. Thanks to its functionality, data sources such as institutional repositories, data repositories, software repositories can be notified of metadata records relative to products (datasets, articles, software, research packages) that are "of interest to them", i.e. metadata records that should be in the data source, or "linked to them", i.e. a scholarly link exists between one of the data source product and the identified product. Notifications are sent only to subscribed data sources, following a subscription and notification pattern, and can be delivered by mail, OAI-PMH end-user interfaces, or, currently under investigation, via push APIs (e.g. SWORD protocol), FTP and ResourceSync. The idea behind the service is to disseminate and advocate the principle that scholarly communication data sources are not a passive component of the scholarly communication ecosystem, but rather active and interactive part of it. They should not consider themselves as thematic silos of products, but rather as hubs of products semantically interlinked with any kinds of research products and, more broadly, up-to-date with the evolving research ecosystem.

## OpenAIRE Advance

This presentation intends not only to focused on what OpenAIRE have achieved building the Open Access infrastructure in Europe, but also outlines the major objectives of the new phase of OpenAIRE to be developed within the OpenAIRE Advance H2020 funded project.

OpenAIRE-Advance will work towards making Open Science the default in Europe, reshaping the scholarly communication system towards openness and transparency serving as a trusted pillar of the European Open Science Cloud (EOSC). In the coming three years OpenAIRE will work along the following lines:

1. Consolidate and boost services: OpenAIRE's Open Science service portfolio will be upgraded to meet end user needs. Through a *set of dashboards* that target all stakeholders involved in the research chain, OpenAIRE will seamlessly connect all research artifacts.
2. Empower the pan-European Open Science Helpdesk: The 34 National Open Access Desks will be empowered to increase their national presence and develop capacities at local level so as to become a *pivotal part of Open Science in national settings*.
3. Strengthen research community uptake of Open Science: Working with three national research infrastructure nodes (Elixir-GR, EPOS-IT, DARIAH-DE) OpenAIRE will build bridges to key communities via an *Open Science-as-a-service* approach.
4. Promote emerging changes in the scientific communication landscape: Building on repositories as the foundation of a globally networked and distributed open science infrastructure, OpenAIRE will support the development of the *next generation repositories* with new functionalities and new technologies.
5. Build a global open science network: Working with partners around the world (Latin America, Japan, US, Canada, Africa) OpenAIRE aims to align policies, practices and services for a truly global and *interoperable scholarly commons*.

## Conclusion

OpenAIRE is a socio-technical network that supports, accelerates and monitors the implementation of Open Science policies, including Open Access to publications and research data relying on a network of content providers such as literature repositories, journals, data repositories and CRIS systems.

OpenAIRE has established itself as a key and sustainable infrastructure for giving access to Open Access publications in Europe, progressively providing access to datasets, software and other research artefacts. From its outset, OpenAIRE has pursued a service-driven design to engage all stakeholders and the current service portfolio (covering all e-Infrastructure layers) targets a variety of users, namely researchers, data providers, funders and the public.

The services for repository managers, research communities and funders highlighted in this presentation are some of the most relevant services within the context of Open Repositories Conference community.

Additionally, this presentation outlines also the major objectives of the new phase of OpenAIRE. OpenAIRE-Advance, starting in January 2018, continues the mission of OpenAIRE to support the Open Access and Open Data mandates in Europe relying on a decentralized network of content providers. By sustaining the current infrastructure, comprised of a human network and technical services, it consolidates its achievements while working to shift the momentum among its communities to Open Science, aiming to be a trusted e-Infrastructure within the realms of the European Open Science Cloud.

## References

[www.openaire.eu](http://www.openaire.eu)