Open Science as-a-Service for Research Communities and Content Providers

Alessia Bardi(✉)

Institute of Information Science and Technologies – CNR, Pisa, Italy
alessia.bardi@isti.cnr.it

Abstract. Open Science is a set of practices of science according to which research activities and the research products they generate should be openly available, under terms that enable their findability, accessibility, re-use and re-distribution. The main effects of the implementation of Open Science principles is to enable responsible, reproducible and transparently assessable research.

For an effective implementation of Open Science principles, a behavioral change in interested stakeholders and new tools for publishing in the scholarly communication ecosystem are required. Open Science publishing calls for the publishing of all types of research artefacts, beyond scientific literature. Today, the scholarly communication ecosystem lacks of tools and research community practices on Open Science publishing. To fill this gap and support a smooth transition towards Open Science, the OpenAIRE initiative is offering two novel services for research communities and content providers (e.g., institutional repositories, data repositories). The final goal is to support the cultural and technological shift towards the Open Science paradigm, from which all the different stakeholders in the research domain and of the society at large can benefit.

Keywords: Open Science · Scholarly communication · Scientific communication · OpenAIRE

1 Open Science Publishing

Open Science is an umbrella term that defines a set of practices of science according to which research activities and any kind of products (e.g. literature, data, methods, software, workflows) they generate should be made available as soon as possible under terms that enable their findability, accessibility, re-use, and re-distribution [1]. Open Science principles are advocated by all scientific and scholarly communication stakeholders [2], who intend Open Science as a means for accelerating research by enhancing transparency and collaboration, and fostering innovation and reproducibility. Scientists and organizations see Open Science as a way to speed up, improve quality, and more effectively reward research activities, while funders and ministries see it as a means to optimize cost of science and leverage innovation [3].

Open Science is still an emerging vision and a way of thinking that calls for behavioral changes in the interested stakeholders and for novel tools in the scholarly

M. Coppola et al. (Eds.): GECON 2018, LNCS 11113, pp. 3–6, 2019.
https://doi.org/10.1007/978-3-030-13342-9_1

communication ecosystem that support the implementation of the Open Science principles.

Open Science publishing means publishing/depositing all products of a research activity under terms that enable their re-use and redistribution. The scholarly communication ecosystem should allow research communities to share and re-use their scientific results by publishing all intermediary and final research artefacts, beyond scientific literature. Artefacts can be research data, software and research methods (e.g. protocols, algorithms, etc.), which should be deposited in repositories for scientific communication (e.g. institutional repositories, data archives, software repositories, CRIS systems), and should be published together with the semantic links between them. To complete the picture, such ecosystem should support publishing of packages of artefacts (e.g. research objects [4], enhanced publications [5]) to allow discovery, evaluation, and reproducibility of science (e.g. workflows or experiments with input datasets). In addition, products of science should be published as soon as possible, i.e. not only once a research activity is concluded, but also while it is ongoing. In other words, scientists should, as much as possible, make their methodologies, thinking and findings available to enable/maximize collaboration and reuse by the research community and the society at large. An effective implementation of Open Science publishing principles would:

- Enable reproducible research: let other users reproduce experiments.
- Enable transparent evaluation of research activities: evaluate findings based on the ability to repeat science, but also on the quality of the individual products of science, i.e. literature, research data, software, and experiments.
- Support re-use of research results: overcome the barriers that SMEs and researchers (especially from developing countries) are facing in accessing publicly funded research.

1.1 Barriers to Open Science Publishing

The existing scientific communication ecosystem lacks tools and practices for engaging research communities at adopting the aforementioned novel Open Science publishing principles. Scientific literature, possibly linked to the underlying research data deposited in a data repository or archive, is still typically seen as the sole, omni-comprehensive unit of scientific dissemination. Publishing other types of artefacts like methods, protocols, workflows and software is not generally a common practice, i.e. no repositories of reference, no persistent identifiers, no citation practices and, therefore, no scientific reward. Regarding the publishing of packages of artefacts, solutions exists in discipline-specific communities (ISA Research Object for systems biology, MINSEQE and MIAMI guidelines for functional genomics) but, in general, research packages are not considered as first-class citizens of the scholarly communication domain.

The scientific ecosystem is also not able to keep a complete and up-to-date record of relationships between research products. For example, publication, data, software repositories and publishers do not keep bi-lateral links between their products, and the links they keep are not in-sync with the updates of the products (e.g. links to new versions of the data, obsolete links).

2 Open Science Publishing as-as-Service

OpenAIRE (Open Access Infrastructure for Research in Europe) [6] is an initiative funded by the European Commission that fosters and enables the adoption of Open Science publishing principles. Its aim is to provide Research Infrastructures (RIs) with the services required to bridge the research life-cycle they support - where scientists produce research products - with the scholarly communication infrastructure - where scientists publish research products - in such a way science is openly accessible, reusable, reproducible, and transparently assessable. OpenAIRE is working closely with existing RIs to (i) provide services for compensating the lack of Open Science publishing solutions and (ii) provide the support required by RIs to upgrade existing solutions to meet Open Science publishing needs (e.g. technical guidelines, best practices, Open Access mandates). To this aim, OpenAIRE is juxtaposing to its "Catch-all repository" Zenodo [7], two new services implementing the concept of "Open Science as a Service" (OSaaS), i.e. services and tools implementing Open Science principles are provided on-demand, transparently with respect to the underlying technical infrastructure. Tools are accessible through either a thin client interface, such as a web browser, or an application program interface. Services are designed to be usable by different disciplines and providers, each with different practices and maturity levels, so as to favor a shift towards a uniform cross-community and cross-content provider scientific communication ecosystem.

The *Research Community Dashboard* will serve research communities and their RIs to publish research artefacts, packages, and links, and to monitor their research impact. Scientists and infrastructural services populate and access an information space of interlinked objects dedicated to them, through which they can share any kind of products in their community, maximize re-use and reproducibility of science, and outreach the scholarly communication at large.

The *Catch-All Broker Service* will engage and mobilize content providers (institutional repositories, data archives, etc.) and serve them with services that help them in keeping their collections of scholarly records up-to-date. 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.

In order to facilitate the cultural and technological shift towards common Open Science publishing practices, the technological efforts are complemented with networking activities that will strengthen the emerging Open Science social environment, aligning practices and mechanisms for open accessibility, transparent evaluation and reproducibility.

3 Conclusion

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 and interlinking (for "discovery" and "transparent evaluation") and re-using (for "reproducibility") all research artefacts produced during a research activity, e.g. literature, research data, methods, software, workflows, protocols.

OpenAIRE fosters transparent evaluation of results and facilitates reproducibility of science for research communities by enabling a scientific communication ecosystem where artefacts, packages of artefacts, and links between them can be exchanged across communities and across content providers. To this aim, OpenAIRE introduces and implements the concept of Open Science as a Service on top of the existing OpenAIRE infrastructure, by delivering services in support of Open Science publishing.

The final goal is to support the cultural and technological shift towards the Open Science paradigm, from which all the different stakeholders of the research domain and of the society at large can benefit in several ways, such as:

- Funders can optimize costs of research activities;
- Researchers can be more effectively credited and rewarded for all the research products they produce, not only for the scientific literature;
- Researcher's activities are reproducible and trustable;
- The open accessibility to all products enables transparent evaluation and support responsible research: SMEs, citizens and researchers worldwide, also from developing countries, can easily access research results, which, in Open Science settings, are not locked behind a paywall.

References

1. FOSTER Open Science Definition. https://www.fosteropenscience.eu/foster-taxonomy/open-science-definition. Accessed 16 Sept 2018
2. European Commission: Validation of the results of the public consultation on Science 2.0: Science in Transition [report]. Brussels: European Commission, Directorate-General for Research and Innovation (2015). http://ec.europa.eu/research/consultations/science-2.0/science_2_0_final_report.pdf
3. European Commission's Directorate-General for Research & Innovation (RTD): Open Innovation, Open Science and Open to the World (2016). https://ec.europa.eu/digital-single-market/en/news/open-innovation-open-science-open-world-vision-europe
4. Bechhofer, S., et al.: Why linked data is not enough for scientists. Future Gener. Comput. Syst. **29**(2), 599–611 (2013). ISSN 0167-739X. https://doi.org/10.1016/j.future.2011.08.004
5. Bardi, A., Manghi, P.: Enhanced publications: data models and information systems. LIBER Q. **23**(4), 240–273 (2014). http://doi.org/10.18352/lq.8445
6. Manghi, P., et al.: OpenAIREPlus: the European scholarly communication data infrastructure. D-Lib Mag. **18**(9), 1 (2012). https://doi.org/10.1045/september2012-manghi
7. Zenodo. https://zenodo.org/. Accessed 16 Sept 2018