

Relevance of EOSC and FAIR in the Realm of Open Science and Phases of Implementing the EOSC

Michel Schouppe
European Commission (Directorate-General for Research and Innovation),
Brussels, Belgium

Michel.Schouppe@ec.europa.eu

Jean-Claude Burgelman
European Commission (Directorate-General for Research and Innovation),
Brussels, Belgium

Jean-Claude.Burgelman@ec.europa.eu

Abstract. In the realm of its Open Science strategy, the European Commission (EC) has announced its intentions to promote open access to research data resulting from publically funded research under Horizon 2020 to increase the circulation and exploitation of knowledge. Openness should become the default case respecting constraints to privacy, national security and intellectual property rights. For making open science happen, two major pillars need to be addressed. Data needs to be FAIR (Findable, Accessible, Interoperable, Reusable) to overcome the fragmentation and inefficiencies and thus to increase data reuse effectively. Commissioner Moedas made clear that "Europe's final transition must be one from fragmented data sets to an integrated European Open Science Cloud (EOSC)". By 2020, all European researchers need to be able depositing, accessing and analyzing European scientific data through the EOSC. EOSC and FAIR research data are closely related and the EC is working out phases for implementing EOSC and action plans to make data FAIR.

Keywords: Open Science, European Open Science Cloud (EOSC), FAIR data management, open data sharing, research infrastructures, e-infrastructures, cloud technologies.

1 Introduction

In a fast moving and increasingly interdependent world, many social, economic and sustainability issues faced by countries and citizens are cross-border and interdisciplinary in nature. Grand societal challenges like climate change, health and energy are intertwined with local and regional issues. Addressing such complex challenges at various scales requires strengthening collaboration across borders and disciplines based on new shared approaches to generate and diffuse scientific knowledge (the latter being increasingly data driven by default).

In 2013, the G8 science ministers recognized the pivotal role of science to secure present and future sustainable growth [7]. Proposed areas for collaboration included global challenges, global research infrastructure, open scientific research data, and increasing access to the peer-reviewed, published results of scientific research.

At the Hangzhou Summit in 2016, the G20 Leaders resolved to promote open science and to facilitate appropriate access to public funded research results through implementation of Findable, Accessible, Interoperable and Reusable (FAIR) principles [8].

Open Science provides a fundamental transformation of the whole research lifecycle to make it more open and accessible, more credible with increased integrity, more reliable and transparent, more efficient and collaborative,

more responsive to societal challenges and closer to citizens. The ongoing transition to open science brings new opportunities to accelerate fundamental research, cross-disciplinary research, cutting edge innovation and co-designed pathways towards sustainability.

Since about five years, the European Commission – at the instigation of Carlos Moedas, the EU Commissioner for Research, Science and Innovation – has placed the concepts of Open Science, Open Innovation and Open to the World at the center of the European policy for research and innovation [6]. Broad consultations and engagement activities across Europe since 2014 have resulted into eight Open Science policy actions, including the following two pillars:

- to develop research infrastructures for Open Science and a common European framework to improve data storage, access, analysis, re-use and governance of research data;
- to mainstream and further promote open access to research data (and metadata), as well as data management practices aiming at making these data Findable, Accessible, Interoperable and Re-usable (FAIR).

2 The European Open Science Cloud

As a general trend, the quantity of data produced by or susceptible to be injected in research life cycles is growing exponentially. With the internet of everything, all reality tends to become a data point and therefore fit for research. Exploiting this trend requires levelling up capacities to discover, access, and process research data

and information of different types, volumes and origins. The challenge becomes more acute when investigating the intersection of multiple fields, disciplines and cultures. This emphasizes an increasing dependency of science and innovation on new Information and Communication Technologies (ICT) including technologies for data analytics and processing on the cloud. It also highlights the pressing need to better address cross-domain interoperability.

As a matter of fact, too many datasets and publications from publicly funded research still remain beyond the reach of most scientists, even of scientists belonging to the same discipline. Obstacles are diverse ranging from a huge variety of individual (or community) data handlings to the low level of resources (e.g. time, budget, skills) invested in data annotation and preservation

The European Open Science Cloud (EOSC) initiative has been proposed in 2016 by the European Commission as part of the European Cloud Initiative to build a competitive data and knowledge economy in Europe [1]. The EOSC initiative is a means to give a strong push in Europe towards the FAIR management of research data and to ensure that European researchers and professionals in science and technology reap the full benefits of the data-driven science.

The ambition is to set up a virtual environment which federates the handling of research data by existing scientific data infrastructures and provides a one-stop-shop (from the point of view of the researcher) for seamless access to services supporting the research community along the data life cycle. Data providers will benefit from services for data annotation, storage, curation and long term preservation on trusted repositories. Data users will be empowered to discover, access, re-use, combine and analyze research data. Amongst its future service offering, the EOSC would also connect to underpinning high-capacity cloud solutions with super-computing capacity.

The EOSC will be discipline-agnostic. As such, the initiative represents a collective venture of which the implementation must remain inclusive of all stakeholders including research communities, research institutions, research infrastructures, research funding bodies and policy makers.

Following the example of the internet, the EOSC will be governed by rules and procedures which are not physically related to a server or institutions. EOSC is therefore sometimes associated to the internet of FAIR research data.

Challenges are multi-dimensional: scientific, technical, cultural and governance-related. Europe's decision to develop the EOSC reflects the willingness to "embrace change, try new things and be willing to take risks to keep European research and innovation at the forefront of modernity and economic growth"[5].

3 The FAIR data principles and the European FAIR data action plan

Varying data management practices by different research

communities and heterogeneous access policies by different research data infrastructures generate complex ecosystems of poorly interoperable data infrastructures. Resulting data silos slow down the circulation of knowledge and prevent cross-fertilization of interdisciplinary research. This kind of situation occurs in Europe and other regions of the world.

The FAIR data principles were first introduced in 2014 and published in 2016 [9] to define a minimal set of community-agreed guiding principles allowing both humans and machines to Find, Access, Interoperate and Re-use data. The initial intention was to support research data providers to follow good data management and data stewardship practices. The principles are being progressively extended to other digital objects (including algorithms, models, and workflows) produced by researchers, as well as to data managed by the public sector.

Policy-related action to support the demand for FAIR data has started in Europe. Leading by example, the European Commission has introduced the obligation of FAIR Data Management Plans (DMPs) in the context of Horizon 2020, the EU Framework Programme for research and innovation in the period 2014-2020. Horizon Europe – the successor of Horizon 2020 after 2020 – is being conceived to support Open Science at all levels [3].

At the same time, the EU Member States are consistently working on open access to and preservation of scientific information. Overall, open access and open science are also consistently included in the national research agendas.

The next challenge will be to turn the FAIR data principles into practice by all European stakeholders across all EU Member States. A FAIR Data Action Plan [2] will be released in November including actions until 2020 to operationalize the FAIR data principles.

4 The EOSC implementation roadmap

An extensive consultation with scientific and institutional stakeholders in Europe took place in 2016 and 2017. This consultation confirmed the demand for a fit for purpose pan-European federation of research data infrastructures, with a view to moving from the current fragmentation to a situation where data is easy to store, find, share and re-use.

The summary outcome of the consultation was presented in March 2018 by the European Commission in the form of a roadmap for implementing the EOSC [4]. This roadmap was broadly welcomed by the Council of the European Union on the occasion of Competitiveness Council on 18 May 2018. Implementation will follow a two-stage approach for the roll-out of the EOSC. The initial phase will run until 2020 and be essentially implemented through a batch of Horizon 2020 projects. The start of the second phase after 2020 will depend on an in-depth evaluation of the initial phase by the European Commission and the Member States. The roadmap till 2020 includes twelve action lines related to (i) the EOSC architecture, (ii) the FAIR data

principles, (iii) services and datasets accessible via the EOSC, (iv) rules of participation and (v) the governance structure of the EOSC.

Before the end of the year, the governance structure will be formally set up as an agreement of the Member States at the ministerial level. A prototype "EOSC Portal" will also be launched and will evolve to become the future universal entry point to the EOSC service and data gateways. The EOSC platform will progressively expand its catalogues of data and services following a continued survey of data infrastructures in Europe interested in taking part in the development phase of the EOSC. Following a recent open consultation, a compliance framework including rules of participation to the EOSC will be refined to take into consideration the diversity of EOSC providers and users. These rules will guide future operationalization of day-to-day transactions across the EOSC federation and strengthen requirements for FAIR-accredited/certified data infrastructures in Europe.

5 Conclusions

2018 is a landmark year for the EOSC. Consultation, engagement and pilot demonstration activities conducted during the last years are now converging to kick-off the initial phase of implementation.

The EOSC typifies an Open Science accelerator aiming at promoting data Commons and developing a FAIR ecosystem in Europe. Implementation will follow a step-wise approach. While, the EOSC will foremost serve the research community in Europe, it will then expand further to the broader community of users including European public authorities, SMEs and citizens.

In that respect EOSC is Europe's answer - as a public good - to the transition to open science.

While EOSC has a pan-European focus, its scope is global. The initiative is framed to become a major EU contribution to a global playing field regarding open and FAIR research data and related services. Over time, the EOSC will reach out beyond its initial core of European infrastructures and gradually widen to interested networks and nodes from global research partners.

Acknowledgments. This paper draws largely on observations and recommendations by the Open Science Policy Platform of the European Commission, the High Level Expert Group on the European Open Science Cloud, the European Commission Expert Group on FAIR data and the group of Horizon 2020 projects involved in the Implementation Roadmap for the European Open Science Cloud [4].

Disclaimer. The views expressed in this paper do not necessarily represent the opinions of the European Commission or any other body; nor do they commit the Commission to implement them.

References

- [1] "European Cloud Initiative - Building a competitive data and knowledge economy in Europe", Communication of the European Commission to the European Parliament, the Council, the European Economic and social Committee of the Regions, COM(2016) 178 final, 19 April 2016
- [2] Hodson S., Jones S., Collins S., Genova F., Harrower N., Mietchen D., Petrauskaitė R., Wittenburg P., FAIR Data Action Plan: Interim recommendations and actions from the European Commission Expert Group on FAIR data, 2018, <https://doi.org/10.5281/zenodo.1285290>
- [3] Lamy P., Bruder Müller M., Ferguson M., Friis L., Garmendia C., Gray I., Gulliksen J., Kulmala H., Maher N., Plents Fagundes M., Wozniak L. A., Zic Fuchs M. "Investing in the European future we want", Report of the independent High Level Group on maximising the impact of EU Research & Innovation Programmes, European Commission publication, © European Union, 2017, ISBN 978-92-79-70571-7, doi:10.2777/477357
- [4] "Implementation Roadmap for the European Open Science Cloud", Commission Staff Working Document, SWD(2018) 83 final, 14 March 2018
- [5] Moedas C., Next Steps in Tackling Societal Challenges, Extract from the statement by the European Commissioner for Research, Science and Innovation, Lund, Sweden, 4 December 2015.
- [6] "Open innovation, open science, open to the world – A vision for Europe", European Commission publication, © European Union, 2016, ISBN 978-92-79-57346-0, doi:10.2777/061652
- [7] "Open scientific research data should be easily discoverable, accessible, assessable, intelligible, useable, and wherever possible interoperable to specific quality standards." G8 Science Ministers Statement, London UK, 12 June 2013, <https://www.gov.uk/government/news/g8-science-ministers-statement>.
- [8] "We support appropriate efforts to promote open science and facilitate appropriate access to publicly funded research results on findable, accessible, interoperable and reusable (FAIR) principles." G20 Leaders' Communiqué, Hangzhou Summit, China, 5 September 2016, http://europa.eu/rapid/press-release_STATEMENT-16-2967_en.htm
- [9] Wilkinson M. D., Dumontier M., Aalbersberg I. J., Appleton G., Axton M., Baak A., Blomberg N., Boiten J-W., da Silva Santos L. B., Bourne P. E., Bouwman J., Brookes A. J., Clark T.,

Crosas M., Dillo I., Dumon O., Edmunds S.,
Evelo C. T., Finkers R., Gonzalez-Beltran A.,
Gray A. J. G., Groth P., Goble C., Grethe J. S.,
Heringa J., 't Hoen P. A. C., Hooft R., Kuhn T.,
Kok R., Kok J., Lusher S. J., Martone M. E.,
Mons A., Packer A. L., Persson B., Rocca-Serra
P., Roos M., van Schaik R., Sansone S-A.,
Schultes E., Sengstag T., Slater T., Strawn G.,
Swertz M. A., Thompson M., van der Lei J.,
van Mulligen E., Velterop J., Waagmeester A.,
Wittenburg P., Wolstencroft K., Zhao J. Mons
B.: The FAIR Guiding Principles for scientific
data management and stewardship. In:
Scientific Data volume 3, Article number:
160018 (2016). doi:10.1038/sdata.2016.18, 15
March 2016