SCIRES it
SCIentific RESearch and Information Technology
Ricerca Scientifica e Tecnologie dell'Informazione
Vol 10, Special Issue (2020), 53-58
e-ISSN 2239-4303, DOI 10.2423/i22394303v10Sp53

CASPUR-CIBER Publishing, http://www.sciresit.it

OPEN SCIENCE, OPEN RESEARCH DATA AND THE ROLE OF IOSSG

*Paola Gargiulo**

* Coordinator of IOSSG, Italy

Abstract

The paper illustrates the objectives and the on-going activities of the Italian Open Science Support Group (IOSSG) whose main goal is to promote the development and the dissemination of Open Science in Italy. It provides support and practical tools to researchers in the different steps of research process with reference to data collection, management, legal protection, access, archiving, preservation within the frame of the European Open Science Cloud (EOSC) initiative.

Keywords

Open Science, Open Research Data, Open Science Support, Research Data Management, European Open Science Cloud

1. *Introduction*

The focus of this paper is to illustrate the objectives and the activities of the Italian Open Science Support (IOSSG). It is convenient to give a brief of overview of the context to better understand the scope and the goals of IOSSG.

The European Commission – Directorate General Connect illustrates its vision of Digital Science and its integration with the funding programme Horizon 2020 (2014-2020) in a concept paper published in 2013, entitled "Digital Science in Horizon 2020". The paper states:

Digital science relies on the combined effects of technological development and cultural change towards collaboration and openness in research. Digital science makes scientific processes more efficient, transparent and effective by new tools for scientific collaboration, experiments and analysis and by making scientific knowledge more easily accessible. At the same time, Digital science enables emergence of new scientific practices, disciplines and paradigms to respond to the new challenges through global distributed collaborations where citizens and society participate as contributors and direct beneficiaries of scientific knowledge. (Digital Science in Horizon 2020, 2013).

Collaboration, openess, efficiency, trasparency, accessibility, reliability, new scientific practices, disciplines, citizen science, global and partecipatory response to societal challenges, well define the features of Science nowdays.

The cumulative and shared process of knowledge to the benefit of humankind is undergoing a new revolution due to the advent of digital technology.

Science has gradually become digital, more data intensive; the change affects all disciplines and all research cycle steps, opens new horizons, requires new competences, new infrastructures, new tools.

Furthermore digital technologies remove barriers, obstacles to openess and offer great advantages that paradoxically are not embraced due to cultural resistance, vested interests and questionable evaluation and assessment metrics.

There is an urgent need of a cultural change to create, store, share and deliver research outputs based on openess, collaboration and sharing. In this respect the European Union, national and international funding agencies, few national governaments, research perfoming organizations, research libraries leagues, Open Science (OS) activists are in the forefront to make this shift.

The main objective is to favor scientific progress and knowledge, innovation, to respond to societal challenges (environment, health, social injustice etc) to avoid duplications, to facilitate reuse and replicability and to ensure research integrity and quality through trasparency.

2. *Open Science: Research Data Management*

There are several definitions of OS in the literature, in this paper we refer to it as making

accessible and understandable the results of scientific research (data and publications). This includes all the activities, tools, methods related to research process from its early stage to its end (open research data, open methodologies, open notebooks, open source sw, open peer reviewing, open access publications, open educational resources).

The focus is particularly on the areas of interest of the Italian Open Science Support Group (from now on IOSSG): open research data collection, data management plan, data fairfication (making data compliant to the FAIR principles)[1], legal issues, archiving, preservation of data.

According the EU research funding programme Horizon 2020[2], digital research data is information in digital form, in particular facts or numbers, collected to be examined and used as a basis for reasoning, discussion or calculation; this includes statistics, results of experiments, measurements, observations resulting from fieldwork, survey results, interview recordings and images.

All data generated or collected digitally contemplate a series of activities to be carried out before, during and after the research is completed, in order to make it findable, accessible, interoperable, retrievable and reusable.

The need of making research data accessible and reusable has been perceived by several scientific communities and research funders. The European Commission in its capacity as research funder and policy maker has mandated open access to all publications (since 2014) and to research data (since 2017) resulting from Eu funded research projects in Horizon 2020.

The funding programme mandates: the deposit of underlying data necessary for validation of research articles in a data repository; the sharing of data (according to the principle as open as possible, as close as necessary by adopting open licenses (CC0 or CC BY) at the time of publication or after an embargo period. It also requires to deposit all other research data generated during the research in a data repository.

Furthermore a Data Management Plan (DMP) is required to describe how research data is handled, organized, licensed and made openly available to the public, and how it will be preserved after the project is completed.

Motivations must be provided in the DMP when versions or parts of the project research data cannot be openly shared due to third-party copyright issues, confidentiality or personal data protection requirements, or when open dissemination could jeopardize the project achievements.

On one hand, this approach, shared also by other research funders, is a concrete attempt to put open science in practice, a way to push forward OS from a funder's perspective, on the other hand researchers and more often their institutions are unprepared and unequipped to meet these requirements.

Before explaining the role of IOSSG, it is worth contextualizing the creation, management, storage, access of research data, and their digital-infrastructures and research infrastructure services in the larger frame of the European Open Science Cloud (from now on EOSC) initiative.

3. EOSC initiative

In the digital era, the constant and exponential production of digital data is globally affecting many aspects of our lives; there are many social, economical, legal, political, cultural implications in data driven societies which are currently investigated in appropriate venues.

Limiting our observation to the deluge of data generated by research, though unevenly distributed among the different disciplines, it is noticeable the urgency for a worldwide virtual environment to store, manage, analyze, retrieve and re-use data. (Covid-19 is a blatant example).

EOSC is an initiative that took shape in 2015 promoted by the EU Commission on Research, Science and Innovation, aimed to create an environment based on the federation of existing and emerging data infrastructure. A federated, globally accessible environment where researchers, innovators, companies and citizen can publish, find and re-use each other's data and tools for research, innovation and educational purposes under well-defined and trusted conditions.

The objective is to provide an environment offering free, open services for data storage, management, analysis and re-use across

[1] See the FAIR principles (Findable, Accessible, Interoperable, Reusable) set by the Force 11 Group.
[2] European Commission. Horizon 2020 online Manual. Open Access.

disciplines. EOSC's goal is to built on existing and emerging horizontal and thematic data infrastructures, bridging current fragmentation and ad-hoc solutions.

EOSC intends to add value (scale, data-driven science, inter-disciplinarity, faster innovation) and leverage past infrastructure investments.

It is structured in the 3 different layers: governance, service and infrastructure. The first layer is related to governance, policies, rules, trasparency, trust, the second layer to services (research support, legal and ethical issues, data management plan, data stewardship statistics, analytics, etc), the third one to the technological infrastructure and related issues (access rights, accessibility, data storage, data archiving, back-up, data manipulation, discovery strategies, catalogue etc.).

EOSC was officially launched in Vienna in November 2018 and is going to be established as a legal entity by the end of 2020.

Currently several ongoing Eu funded projects are in place to implement the roadmap to build this complex federated architecture which involves research communities, research and digital infrastructures, data repositories and services, research perfoming organizations, governaments.

4. IOSSG and the Italian academic context

The Italian Open Access Support Group has been established in 2016 at a time when the Horizon 2020 Open access mandate to research data and to draft a data management plan was limited to some pilot subject areas and the EOSC initiative was still at an early stage.

Few experts with different skills from universities research services (research project support, ICT, digital libraries, Open Science, legal issues, communication) gather together with the aim in the first place to share their experiences in giving support to researchers in their compliance to Horizon 2020 open research data requirements.

Everyone was experiencing the difficult task of giving assistance on issues related research data management (writing a plan, metadata assignments, selection of a data repository, legal issues, personal data protection just to mention few) to provide answers, suggestions, to dispel doubts.

This effort was rather challenging in absence of plans, policy, guidelines, strategies at national level and at institutional level. A lot of questions were arising and new solutions, new skills were required.

IOSSG members committed themselves to raise awareness and take action on these issues at their respective institutions and to establish informal contact with the Italian node of Research Data Alliance and with ICDI - Italian Computing and Data Infrastructure. This group consists of representatives of the main Italian Research Infrastructures and Digital Infrastructures. Their aim is to promote synergies at national level in order to enhance the Italian participation in the European Digital Science initiatives including EOSC and the European Data Infrastructure (EDI) and HPC (High Performing Computing).

A brief description of the local context helps to understand the reasons underlying the formation of IOSSG.

Within the universities, the knowledge built on research data handling by few data intensive research communities, is fragmentary and rarely focused on sharing, reuse, data preservation. Publishing results in peer reviewed journals is a priority and consequently these other issues have been often neglected up to the day when research funders began to mandate Open Access to research data and the implementation of a DMP.

Furthermore in 2016 Italian universities were lacking institutional policies, clear governance, or strategy to support or to redirect the different requests related to open research data management issues. Given the novelty of digital research data handling which is a different issue from digital data analysis, universities are short of the expertise and competence represented by new professionals such as Data Management Specialist, Data Support Officer, Data Repository Manager, Research Data Management Project Officer, Research Data Management Service Officer, Digital Research Librarian, Data Stewards, Data Curators, etc.

One must admit that handling digital research data is a huge task. Some countries (UK, Netherlands) have started to embark on this project several years ago at national level.

In Italy research data handling and practing Open Science are not yet considered a priority by the university system with very few exceptions, (the data intensive research centres context is quite different), and just recently the national governament has made its first steps to put Open Science on its agenda.

Given this situation, the founders of IOSSG felt the urge to put their efforts together, promote and develop Open Science policies in their institutions and respond to researchers' requests by preparing materials, tools that really suit their needs.

The group is informal, on a voluntary basis, has no istitutional endorsement or funding, keeps contacts with other similar groups in Europe.[3] It has adopted few operating rules: online quarterly plenary meetings, yearly assignment of 1 or 2 tasks, 4 working subgroups, open availability of the outcome (CC BY license) on the IOSSG website.

IOSSG is aware of the role and the importance of the EOSC initiative, for this reason signed the Statement in support of EOSC in 2017 and is part of the EOSC Coalition of doers.

IOSSG developed its work plan taking also into account recommendations, guidelines, tutorials, toolkits, policies, training courses etc made available by several Eu funded projects (OpenAIRE, FOSTER) and particularly by the Eu funded project LEARN.

5. IOSSG outcomes

As it was mentioned before, the Group has formed 4 different working subgroups: 1 devoted to governance issues, 2 to services and 1 to data infrastructure.

Each working group has a coordinator and yearly selects a project to work on remotely by using the most appropriate open source technological media and collaborative writing tools. The working document is deposited in shared storage platform, and the final version, once approved by all members, is made available open (CC BY license) on the IOSSG website.

Three or four times a year, plenary sessions take place, each group reports on the advancement of their works, shares news, information about events attended. During the rest of the year information, exchange of opinions, comments are shared through a mailing list and for brief news through the IOSSG twitter account.

Since 2017 the following output has been produced and is openly available:

a policy model on research data management, (Modello di Policy sulla gestione dei dati, 2017);

a research data management plan checklist (Griglia per l'elaborazione dei dati della ricerca, 2017);

a set of F.A.Q on databases and intellectual property rights. (FAQ in materia di banche dati e proprietà intellettuale, 2018);

a study for the creation of a single, visible access point within a research performing organization that includes cross disciplinary skills to support researchers. (Single Point of Entry. Uno studio per i servizi alla ricerca di Ateneo nel contesto di EOSC, released in early 2019);

contribution to the Italian translation of the Open Science Training Handbook produced by the EU funded project FOSTER. (Manuale per i formatori di Scienza aperta, 2019).

Currently a step by step tutorial on how to write a data management plan is in progress. As well as two more projects are under the way: a collection of useful informative resources on Open Science and research data management geared to different targets within the university (researchers, doctoral students, technical and administrative staff) and a concise and appealing presentation on relevant and crucial issues related to open research data addressed to researchers in an infographic format.

A comparative analysis of data repository solutions and their compliance to the FAIR principles is a task of the Data infrastructure group which has not yet been finalized and hopefully will be ended early next year.

All IOSSG works will also be available on the Portal of Open Science which will be released in the next few months by CNR-ISTI, a site collecting news, information, a catalogue of resources and much more addressed to all Italian research communities and stakeholders.

Among these different useful and practical documents, the policy model and the study on the single point of entry merit a special mention. The former deals with the different roles, responsabilities, tasks to be carried out in handling research data within a university. The document has been revised, discussed and enriched by the University of Milan, the first Italian university to adopt it; subsequently University of Padua and University of Bologna drafted their own version of the policy.

[3] The founding members are experts from: the Italian National Desk- Open AIRE, University of Milan, University Ca' Foscari Venezia, Polytechnics of Milan, University of Turin, University of Bologna, University of Trento, University of Trieste, University of Parma, University of Padua, University of Vienna, the Italian company 4Science. On a personal level Simone Sacchi- IUE and Francesca De Donato- CNR-Pisa take part in the group.

The model was submitted to the CRUI Library Commission Open Access Working Group (some members of IOSSG were also members the CRUI group) and luckly resulted in the approval of a policy model which any Italian university can adopt and personalize.

The document on the single point of entry is an articulated study inspired by the EOSC layers previously mentioned. The study presents a proposal for the creation of a single, visible access point within a research performing organization that includes cross disciplinary skills to support researchers, as designed in EOSC service layer.

The underlying idea is that setting policies, services, infrastructures in local context requires the definition of new intersecting processes related to the digital workflow of research. These processes are distributed across the whole research performing organization (university or research institute). A single point of entry at local level gathers this cross knowledge, develops and provides as set of competencies and skills. After a brief description of the EOSC initiative, the document envisages how these three levels can be implemented locally. The objective is to create links among the scattered competence and activities and build new skills and new support services to manage data.

6. Conclusions

IOSSG is committed to support stakeholders in facing the challenges data driven science poses to the academic and research institutions.

Within its limited capacity it is contribuing to empower researchers and institutions with useful tools and skills concerning research data management, governance and policy issues.

It is confident that the work done is a valuable step in fostering Open Science in practice and is open to collaboration with other groups and organizations that share the same practical and goal oriented approach.

REFERENCES

CRUI. Commissione Biblioteche - Gruppo Open Access (2017). *Modello di policy di ateneo relativa ai dati della ricerca*. Retrieved from http://www.ceit-otranto.it/images/documents/crui-oa-modello-policy-ateneo-dati-ricerca-2017.pdf

European Commission. European Open Science Cloud (EOSC). Retrieved from https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud

European Commission. Directorate General - Connect. (2013). Digital Science in Horizon 2020. Concept paper of a vision of Digital Science and its integration with the funding programme Horizon 2020. Retrieved from https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=2124

European Commission. Open Access. In *Horizon 2020 online Manua*l. Retrieved from https://tinyurl.com/wv4gfok

Force11 Group (2016). *The FAIR data principles*. Retrieved from https://www.force11.org/group/fairgroup/fairprinciples

FOSTER (2018). *Il Manuale per i Formatori di Scienza Aperta* (Italian translation, 2019). Retrieved from https://book.fosteropenscience.eu/it/

ICDI- Italian Computing and Data Infrastructure. Retrieved from https://www.icdi.it/it/

IOSSG- Italian Open Science Suppot Group. Retrieved from https://sites.google.com/view/iossg/

IOSSG (2017). *Griglia per l'elaborazione del piano di gestione della ricerca*. Retrieved from https://tinyurl.com/y52eouaj

IOSSG (2017). *Modello di policy sulla gestione dei dati della ricerca*. Retrieved from https://tinyurl.com/y3ozxfue

IOSSG (2018). *FAQ in materia di banche dati e diritti di proprietà intellettuale*. Retrieved from https://zenodo.org/record/2875400#.X2Jl5xAzbIV

IOSSG (2019). *Single Point of Entry. Uno studio per i servizi alla ricerca di Ateneo nel contesto di EOSC. Attività, competenze e ruoli*. Retrieved from https://zenodo.org/record/2602139#.XKHqZ9hS_IU

LERU - League of European Research Universities (2013). *LERU Roadmap to Research Data.* Retrieved from https://www.leru.org/publications/leru-roadmap-for-research-data

LERU (2018). *Open Science and its role in universities: a roadmap for cultural change*. Retrieved from https://www.leru.org/publications/open-science-and-its-role-in-universities-a-roadmap-for-cultural-change

LEARN. (EU funded project to create resources to help Research Performing Institutions to manage their research data) Retrieved from http://learn-rdm.eu/en/about/

OpenAIRE. (EU funded project collecting and making access to OA publications and data resulting from EU and nationally funded research projects). Retrieved from https://www.openaire.eu/

Research Data Alliance- RDA- Italy. Retrieved from https://rd-alliance.org/groups/rda-italy

Università degli Studi di Milano (2017). *Policy sulla gestione dei dati della ricerca*. Retrieved from https://www.unimi.it/sites/default/files/regolamenti/PolicyRDM.pdf

Università degli Studi di Padova (2018). *Policy sulla gestione dei dati della ricerca*. Retrieved from https://www.unipd.it/sites/unipd.it/files/2018/policy%20dati%20ricerca.pdf