

Open Science&Research Data Management – a contribution to online training resources for research institutions

Gabriel NEAGU, Ella Magdalena CIUPERCĂ, Dragoș-Cătălin BARBU, Alexandru STANCIU, Andreea DINU, Sorin CONSTANTINESCU

National Institute for Research & Development in Informatics – ICI Bucharest,
8-10 Maresal Averescu Avenue, 1st District, Bucharest, Romania

gabriel.neagu@ici.ro, ella.ciuperca@ici.ro, dragos.barbu@ici.ro, alexandru.stanciu@ici.ro,
andreea.dinu@ici.ro, sorin.constantinescu@ici.ro

Abstract: *Open Science (OS) and Open Access (OA) to research results (scientific publications, data sets) are widely recognised and supported concepts due to their benefits regarding the quality, integrity and efficiency of the research activity. Research data management (RDM) is considered a key component of research integrity, simplifying the processes of discovery, evaluation, and reuse of research data as FAIR and open data. The European Open Science Cloud (EOSC) provides an integrating and evolving ecosystem to demonstrate the benefits of implementing these concepts for providers and users of various research results and services. Also, the European and National Research & Innovation Programmes play a pivotal role in increasing the awareness of research organisations about the priority of getting prepared to implement these concepts in their research projects financed by public funds. Taking benefits of participating in a Horizon 2020 project dedicated to boosting and guiding the SEE countries contribution to the EOSC, ICI Bucharest developed a training support on OS&RDM for research organizations. The paper presents the content structure of these online training resources, as a component of a more complex training service, including online/F2F courses and practical/hands-on activities, to be provided through the ICI Center for Continuous Training and Professional Development.*

Keywords: Open Science, Open Access, Research Data Management, NI4OS-Europe project, training resources.

1. Introduction

Open science (OS) is the movement to make scientific research, data and dissemination accessible to all levels of an inquiring society (FOSTER, 2018). It is the practice of science in a manner that allows third-parties' collaboration and contribution, where research data, laboratory notes, and other research processes are freely available, under conditions that allow the reuse, redistribution, and reproduction of the research, and of underlying data and methods at its base. *The Horizon Europe Programme Guide of the European Commission* defines the Open science as an approach based on open cooperative work and systematic sharing of

knowledge and tools as early and widely as possible in the process (European Commission, 2024). The main advantages of implementing this concept are related with the potential to increase the quality and efficiency of research, to share the results as accelerator of the advancement of knowledge and innovation, to improve the reusability and reproducibility of results, to support the involvement of all relevant knowledge actors. In its *Recommendation on Open Science*, UNESCO formulates a broader view of this concept: Open science is a set of principles and practices that aim to make scientific research from all fields accessible to everyone for the benefits of scientists and society as a whole. Open science is about making sure not only that scientific knowledge is accessible but also that the production of that knowledge itself is inclusive, equitable and sustainable (UNESCO, 2021). OS encompasses all disciplines and scientific aspects of academic practices, including basic and applied sciences, natural and social sciences, humanities and is based on the following key pillars: open scientific knowledge, open scientific infrastructures, scientific communication, open involvement of societal actors and open dialogue with other knowledge systems. In this vision, the OS contribution is summarized as follows: it amplifies scientific collaborations and the exchange of information for the benefit of science and society; ensures the availability, accessibility and reusability of multilingual scientific knowledge for all; it opens the processes of creation, evaluation and communication of scientific knowledge to societal actors, beyond the traditional scientific community.

Open access (OA) is defined as the practice of providing online access to research results free of charge to end users and without restrictions on use and reuse, beyond the possibility of claiming authorship. In the Berlin Declaration, OA is defined as unrestricted, online access to peer-reviewed, scholarly research papers for reading and productive re-use, not impeded by any financial, organisational, legal or technical barriers. Ideally, the only restriction on use is an obligation to attribute the work to the author (Open Access Initiative, 2003). The *Declaration on Research Assessment (DORA)*, adopted in San Francisco in 2012, included among its recommendations the implementation of OA to both research articles as well as research data and methods used to calculate all metrics (DORA, 2012). In 2015, *Science Europe* established the principles of OA to research publications with the aim to shift to a research publication system in which free access to research publications is guaranteed, and which avoids undue publication barriers, replacing the subscription system with other publication models (Science Europe, 2015). According to European Commission (2017), the main advantages of open access to publications and research data are: improving the quality of research results by taking into account previous results, encouraging collaboration and avoiding duplication of efforts, accelerating innovation, involving citizens and society through transparency of the scientific process. In 2018, *cOAlition S* launched *the Plan S*, a set of principles that ensure open and immediate access to funded research publications. *cOAlition S* is an international consortium of funders and stakeholders who are committed to making full and immediate OA to research publications a reality. According to Plan S, "research funders will mandate that

access to research publications that are generated through research grants that they allocate, must be fully and immediately open and cannot be monetised in any way.” (cOAlition, 2018). For the field of scientific research, OA refers to the possibility for research results in digital format to be accessed and reused with as few restrictions as possible. OA policies aim to provide researchers and the general public with free access to peer-reviewed scientific publications, research data and other research results, in an open and non-discriminatory manner, as quickly as possible during the research process, and to enable use and reuse of these results (European Commission, 2018).

The EU Research Framework Programmes *Horizon 2020* and *Horizon Europe* played a pivotal role in promoting and implementing OS/OA principles and practices. In H2020, the European Commission ran *the pilot ORD (Open Research Data)* on open research data, which aimed to improve and maximize access to research data generated by H2020 projects, as well as their reuse (European Commission, 2017). Costs associated with data management, including the development of data management plans (DMPs), were eligible under H2020 grants. This pilot focused primarily on the data needed to validate the results presented in scientific publications. *The Horizon Europe Program Guide* nominates as mandatory OA practices for all beneficiaries of grants funded by the programme: (i) OA to scientific publications and research data (according to the principle "as open as possible, but as closed as necessary"); (ii) responsible management of research data in accordance with the principles of FAIR (Findable, Accessible, Interoperable, Reusable); (iii) the availability of information about research results or tools needed to validate the conclusions of scientific publications or to validate/reuse research data (European Commission, 2024).

At the national level, the free access to scientific publications and research data is mentioned in the specific Objective 1.2 "Ensuring the transition to open science and facilitating the path to excellence in scientific research" of the National Strategy for Research, Innovation and Smart Specialization 2022-2027, elaborated by the Ministry for Research, Innovation and Digitalization (MCID, 2022a). Also, in the National Plan for Research, Development and Innovation 2022-2027, Open Science is nominated as one of its building principles (MCID, 2022b). The White Book of the transition to OPEN SCIENCE 2023-2030 is a strategy paper on the framework for developing Open Science in Romania, which presents the vision of OS development in Romania by 2030 and targets for achieving the vision, 8 strategic objectives supporting this vision, a list of the involved actors and an index of proposed actions (UEFISCDI, 2023). The Objective 6 - "Developing the Capabilities Required for Open Science Implementation" underlines the importance of developing/ adapting national policies on the skills and competencies of researchers and staff in academic and research institutions related to scientific information and open science, in order to provide education and training as part of the higher education and training system. In this context, increasing awareness and practical knowledge about OS/OA at the level of the

research community should be a priority to prepare its effective participation to the implementation of the above-mentioned vision and, in particular, for successful participation at national and international RDI calls.

ICI Bucharest is a pioneering research institution in OS/OA promotion at both the national and institutional levels (EOSC Focus, 2024). This paper aims at presenting the ICI Bucharest's training resources to support research institutions in achieving the necessary level of understanding of main priorities and related policies, practical activities and necessary skills for implementing these concepts. The framework for elaborating this training support was provided by the H2020 grant 857645 "NI4OS-Europe - National Initiatives for Open Science in SEE countries". The general objective of the grant was to activate the SEE countries contribution to the European Open Science Cloud (EOSC). It should be mentioned that since its launch, EOSC has provided a valuable and powerful environment for practising the implementation and demonstrating the benefits of OS/OA concepts. Currently, the EOSC Association has 236 member organizations from 38 countries.

In the training support a special emphasis is put on Research Data Management. The relevance of this subject in the OS/OA context is argued in Section 2. Section 3 presents the content structure of the proposed training support, which covers 5 main topics. Some conclusions are formulated in the final section.

2. The role of Research Data Management (RDM)

Open research data (ORD) is data from research with open access. According to *the EU Open Data Directive*, ORD are data from research projects that have benefited from public funding, for which open access is ensured according to the principle of "openness by default" and compatibility with the FAIR principles: Findable, Accessible, Interoperable, and Reusable (European Parliament, 2019). A recognized contribution to the maturation of the ORD concept was the document entitled *Concordat on Open Research Data*, developed with the contribution of the UK Open Research Data Forum, under the auspices of UK Research and Innovation (UKRI) - the national funding agency for science and research in Great Britain (UKRI, 2016). Among the 10 principles proposed in the document, 3 refer explicitly to open access to research data:

- Open access to research data is an enabler of high quality research, a facilitator of innovation and safeguards good research practice;
- There are sound reasons why the openness of research data may need to be restricted but any restrictions must be justified and justifiable;
- Open access to research data carries a significant cost, which should be respected by all parties.

Another principle states that proper data management is fundamental to all stages of the research process and should be established from the outset.

Regarding the FAIR principles, their meaning is explained, for example, in (FOSTER Plus, 2018). The FAIR principles are guiding principles in that they describe the qualities necessary for data to be maximally reusable. These principles can apply to data regardless of their public availability, and specifically do not require data to be open. But, considering the objective of reuse, priority is their application for data created by publicly funded research, according to the motto "as open as possible, as closed as necessary". Therefore, FAIR and Open should be considered as complementary concepts.

Research data management (RDM) refers to the organization, storage and preservation of data created during a research project. It covers initial planning, day-to-day processes, archiving and long-term sharing or deletion. RDM is a key component of research integrity, which in turn provides the foundation for public trust in research and the new knowledge it generates. In *the Horizon Europe Programme Guide*, RDM is defined as the process within the research lifecycle that includes the data collection or acquisition, organisation, curation, storage, (long-term) preservation, security, quality assurance, allocation of persistent identifiers (PIDs), provision of metadata in line with disciplinary requirements, licencing, and rules and procedures for sharing of data. RDM is considered an essential element in any project that generates, collects or re-uses data (European Commission, 2024).

Good management of research data should simplify the processes of its discovery, evaluation and reuse. The FAIR principles are therefore considered to be a set of essential (golden) guidelines for RDM. In other words, ensuring data is FAIR is an essential precondition for a proper RDM (Wilkinson et al., 2016).

The general objectives of RDM are:

- compliance with the requirements formulated by the legislation in force, by the funder and by the research organization;
- facilitating data security measures (confidentiality, integrity, availability);
- adequate facilitation of access, collaboration and sharing of data and results.

RDM brings benefits to the researcher, the research organization, and the research community as a whole, such as: ensuring the safe storage of research data, the ability to validate research results, and the ability to reuse and share datasets. In particular, the benefits of sharing and reusing research data refer to:

- encouraging scientific inquiry and debate;
- promoting innovation and potential new uses of data;
- generating new collaborations between data users and their creators;
- maximizing transparency and accountability;

- facilitating the examination and control of research results;
- reducing the cost of duplicating data collections;
- increasing the impact and visibility of the research that created the data and its results;
- the possibility of crediting the researcher as the author of the result obtained;
- providing important resources for education and training.

The exploitation of these benefits is facilitated by ensuring the open character of the data. According to Commission Recommendation (EU) 2018/790 on access to scientific information and its preservation, Art.3, Member States should establish and implement clear policies for the management of research data resulting from publicly funded research, whereby to also ensure that data in this category becomes and remains findable, accessible, interoperable and reusable ("FAIR principles") in a safe and reliable environment, through digital infrastructures, including those grouped in the European Open Science Cloud - EOSC (European Commission, 2018).

3. Topics of OS&RDM training resources

A. Legal and Policy issues

A.1. Legal aspects regarding research data

- Research data reuse: Research data and open data; Timeline of open access initiatives; Steps and recommendations for opening data; Ways to reuse research data; Examples of platforms and resources for research data reuse; Research data and open data;
- Protected rights for research data: Legal aspects regarding the intellectual property of the data; Protection and Licensing of Research Data; Copyright and Intellectual Property; Rights protected in relation to research data: Copyright, Sui Generis Database Right;
- Protection of research data for reuse: How can a protected data set be used?
- Licensing research data for open access: Tools available for license selection; Creative Commons (CC) licenses; Open Data Commons (ODC) license; GitHub Licenses;
- Legal aspects regarding data interoperability: Legal Interoperability of Research Data: Principles and Implementation Guide (https://zenodo.org/records/162241#.Y_hmu3ZByUm).

A.2. Policies at the institutional level for OS&RDM

- Terminology OS&RDM;
- Model Open Science (OS) Policy for Research Performing Organizations (RPO), OpenAIRE Advance project: Rights, Responsibilities and Duties for RPO and for researchers; Open access to publications; Open access to research data; Other OS practices considered; Validity of the Policy;
- Research Data Management (RDM) support capabilities and their role: Politica RDM; Plan de afaceri și sustenabilitate; Planificarea managementului datelor; Managementul activ al datelor; Acces și publicare; Estimarea și evaluarea riscurilor; Conservarea datelor; Consultanță; Participarea la EOSC; Evaluarea FAIR;
- Model Policy for RDM at Research Institutions, based on the analysis of 23 case studies regarding the experience of prestigious universities and research institutions in the development of policies dedicated to RDM (H2020 Grant 654139 ”*LEARN - Leaders Activating Research Networks*”); Detailing the meaning of the document: Preamble; Jurisdiction; Intellectual property rights; Handling research data; Responsibilities, rights, duties for researcher and research organizations; Validity.

A.3. Incentives and rewards in promoting OS&RDM

- Criteria: Scientific quality before quantity, Reproducibility of results, FAIR data, Data organization and archiving, Sharing data and result, Data reuse, Interaction between stakeholders to identify research requirements and collaboratively create results;
- Support actions for fulfilling the criteria;
- OS incentives and rewards for researchers; OS incentives and rewards for institutions;
- Conventional OS metrics; Alternative OS metrics – Altmetrics: chronology of regulatory actions, examples, advantages and challenges; Recommendations for a new generation of metrics;
- Coalition for Advancing Research Assessment (CoARA) and CoARA Agreement.

B. European Open Science Cloud (EOSC)

B.1. EOSC presentation

- overview - EOSC as a crosscutting data space for Research and Innovation; EOSC Association;

- EOSC structure as a compound service for science and society, the role Minimum Viable EOSC (MVE);
- Typology of EOSC resources and European projects dedicated to populate the EOSC ecosystem;
- EOSC Strategic Research & Innovation Agenda and its Multi-Annual Roadmap; Reflection paper - EOSC post 2027.

B.2. Methods of interaction with the EOSC portal

- the role of the EOSC portal;
- target groups of portal users and provided benefits;
- accessing the portal, the user's profile and the associated dashboard;
- resource and service user's perspective: marketplace - access to catalogs by types of resources, the use of search filters, information available for the resources accessed, comparison of resources and services, methods of access, possibilities to contact a supplier;
- resource and service provider's perspective: types of accounts, required descriptive information, on-boarding procedure steps;
- use-cases provided by EOSC users.

B.3. NI4OS-Europe AGORA Catalogue - Southeast European contributions to EOSC resources

- AGORA catalogue dashboard-categories of resources: core services, generic services, thematic services, publications and data repositories;
- examples of resources for generic services: HPC resources, Cloud resources, storage resources, data analysis service;
- examples of thematic services for life sciences, cultural heritage, climate change;
- use case: an open call for AGORA resources: 20 projects from EU and candidate countries, which accessed computing & storage resources and 16 thematic services;
- integration of the AGORA catalogue with the EOSC Portal Catalogue & Marketplace.

C. Open Research Data

C.1. FAIR data: requirements, metrics, FAIR-ification

- FAIR principles and requirements for (meta)data, FAIR data vs Open data, benefits of FAIR principles;

- Overall FAIR framework: Turning FAIR into Reality – Final report and action plan from the EC Expert Group on FAIR Data;
- Data FAIRification - an evolutive process: partly FAIR may be fair enough; generic workflow and necessary competencies;
- FAIR metrics: FAIRsFAIR Data Object Assessment Metrics, RDA FAIR Data Maturity Model, Recommendations on FAIR Metrics for EOSC by Executive Board FAIR Working Group.

C.2. Research data management planning and compatibility with FAIR principles

- Data Management Plan (DMP): objectives, document structure, components of DMP implementation framework, DMP elaboration workflow;
- DMP & FAIR principles: how DMP contributes to FAIR principles fulfilment;
- DMP compatibility with FAIR principles: challenges and solutions, benefits and impact, good practices at institutional and project levels;
- implementation tools for DMP and FAIR principles;
- DMP model adopted in the Horizon Europe programme.

C.3. Support tools in the management of open research data

- Amnesia – OpenAIRE tool for data anonymization (<https://amnesia.openaire.eu/index.html>): objective; steps for loading data; steps for creating generalized hierarchies; steps for anonymization of entered data;
- RoLECT – NI4OS tool for checking compliance with EOSC participation rules of research data (<https://rolect.ni4os.eu/rolect/>): objectives; the workflow; quizzes: general information, service transparency information, Intellectual Property Rights information, Intellectual Property Rights restrictions, multi-licensing Intellectual Property Rights, personal data information, ethics terms information, public sector information; information loading;
- RePol – NI4OS Repository Policy Generator (<https://repol.ni4os.eu/faces/index.xhtml>): objectives; policy generation; types of policies; structure of data repository policy (general info, metadata, data/content access, content, submission of additional data, versioning, privacy, harvesting, curation, preservation), structure of privacy policy (general information, legal aspects, personal data collection, Data Privacy Officer Contact, Supervisor Contact, policy updates); downloading the generated policies.

D. Open Software licensing

- Intellectual property rights: IPRs relevant for OS; Relevance of Software licenses,
- Creative Commons licenses: CC typology; CC licenses for software
- Software licenses: Free, Open source, Proprietary
- Licențe Copyleft/Share-Alike
- Copyright
- Free and Open Source Software (FOSS): the basics; FOSS licenses; FOSS – Rights and Obligations

E. OS&RDM requirements for HE projects

- Experience of Open Access promotion in Horizon 2020: to scientific publications and to research data;
- H2020 results of the OA policy;
- Open science in Horizon Europe: mandatory and recommended OS practices;
- OS provisions in the Horizon Europe - General Model Grant Agreement: OA to scientific publications and RDM;
- OS approach in the Horizon Europe proposal: Part A: Application form, Part B: Project proposal.

4. Conclusions

In the context of the dynamic evolution of the interest for OS/OA, the paper mainly presented the content structure of online training support focused on the priorities for researchers and their organizations to face the requirements generated by this evolution. The support takes advantage of the authors' participation in a Horizon 2020 grant dedicated to National Initiatives for Open Science in SEE countries and their capacity and commitment to continuously update and adapt the content according to the latest developments in the tackled topics. Moreover, as compared with many other offers in this domain, the proposed contribution is packaged in a more complex training service which is going to be provided by ICI Bucharest through its Center for Continuous Training and Professional Development. This service includes courses detailing this support and practical/hands-on related activities.

REFERENCES

cOAlition S (2018) *Plan S - making full & immediate Open Access a reality*. cOAlition S. European Science Foundation, 2018. https://www.coalition-s.org/plan_s_principles/

DORA (2012) Declaration on Research Assessment. *The Annual Meeting of The American Society for Cell Biology (ASCB)*, San Francisco, CA, December 16, 2012. <https://sfdora.org/read/>

EOSC Focus (2024) *Romania's strategic role in advancing Open Science and research data management: Insights from Victor Vevera and Dragoș-Cătălin Barbu of ICI Bucharest*. EOSC Focus Project, 31 Oct. 2024. (<https://eosc.eu/news/2024/10/romaniias-strategic-role-in-advancing-open-science-and-research-data-management-insights-from-victor-vevera-and-dragos-catalin-barbu-of-ici-bucharest/>). [Accessed 10th August 2024]

European Commission (2017) *H2020 Programme - Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020*. https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf [Accessed 10th August 2024]

European Commission (2018) Commission Recommendation (EU) 2018/790 of 25 April 2018 on access to and preservation of scientific information. *Official Journal of the European Union, OJ L 134, 31.5.2018*, pp. 12-18. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32018H0790>. [Accessed 11th August 2024]

European Commission (2024) *Horizon Europe (HORIZON) Programme Guide*. Version 4.1, 01 May 2024. https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/programme-guide_horizon_en.pdf [Accessed 11th August 2024]

European Parliament (2019) DIRECTIVE (EU) 2019/1024 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 20 June 2019 on open data and the re-use of public sector information (recast). *Official Journal of the European Union, L 172/56, 26.6.2019*. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019L1024&from=EN> [Accessed 11th August 2024]

FOSTER Plus (2018) *The Open Science Training Handbook*. H2020 Project 741839 "Fostering the practical implementation of Open Science in Horizon 2020 and beyond". https://open-science-training-handbook.github.io/Open-Science-Training-Handbook_EN/ [Accessed 13th August 2024]

MCID (2022a) *The National Strategy for Research, Innovation and Smart Specialization 2022-2027*. Ministry of Research, Innovation and Digitalization,

Government Decision no 933 from 20 July 2022. https://www.poc.research.gov.ro/uploads/2021-2027/conditie-favorizanta/sncisi_19-iulie.pdf. [Accessed 13th August 2024]

MCID (2022b) *The National Plan for Research, Development and Innovation 2022-2027*. Ministry of Research, Innovation and Digitalization, Government Decision no 1188 from 29 September 2022. <https://www.research.gov.ro/wp-content/uploads/2022/12/hg-aprobare-pncdi-iv.pdf>. [Accessed 13th August 2024]

Open Access Initiative (2003) *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities*. The Max Planck Society, Berlin, 2003. <https://openaccess.mpg.de/Berlin-Declaration> [Accessed 13th August 2024]

Science Europe (2015) *Principles on Open Access to Research Publications*. Science Europe Working Group on Open Access. https://www.scienceurope.org/media/4kxhtct2/se_poa_pos_statement_web_final_20150617.pdf [Accessed 14th August 2024]

UEFISCDI (2023) *The White Book of the transition to OPEN SCIENCE 2023-2030. Strategy paper on the framework for developing Open Science in Romania*. Coordinated by UEFISCDI - Executive Agency for Higher Education, Research, Development and Innovation Funding. https://uefiscdi.gov.ro/resource-863421-brosura-open-science-en_v1-smallest-text.pdf [Accessed 15th August 2024]

UKRI (2016) Concordat on Open Research Data. July 2016. <https://www.ukri.org/wp-content/uploads/2020/10/UKRI-020920-ConcordatonOpenResearchData.pdf>. [Accessed 15th August 2024]

UNESCO (2021) *Recommendation on Open Science*. Adopted by the General Conference of UNESCO at its 41st session, in November 2021. <https://en.unesco.org/science-sustainable-future/open-science/recommendation> [Accessed 15th August 2024]

Wilkinson, M. et al. (2016) The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). doi:10.1038/sdata.2016.18.