PUBLIC (HEALTH) SECTOR AND ACADEMIA

Matej Antol, Michal Růžička, Luděk Matyska, Jiří Marek

OPEN SCIENCE AND EOSC: RESEARCH DATA IN EVERYDAY ACADEMIC LIFE

INFORMATION SOCIETY IN ACADEMIA[1]

"Open Science is a system change allowing for better science through open and collaborative ways of producing and sharing knowledge and data, as early as possible in the research process, and for communicating and sharing results." European Commission [1]

Following many years of preparation and evolution, Open Science is becoming a standard of scientific life[2]. Researchers worldwide experience the effects of opening research and accessing its outcomes in their day-to-day work. The current topics include Open Access[3], increased sharing of research data (FAIR Data = Findability, Accessibility, Interoperability, and Reuse [2,3]), collaboration and other components of the scientific process – from open methodologies, peer review, software and tools to new methods of research quality assessment [4]. At the same time, the findings of publicly funded research are becoming interconnected with society (Citizen Science). Open Science represents a new, modern way to implement research and open access to scientific knowledge through digital technologies and tools enabling advanced cooperation.

To be "an internationally recognised research university with excellent higher education approach, there is a need for academic institutions to set trends in fulfilling all the roles of a university."[4] As such, these progressive institutions cannot stay aloof from the trends transforming global science; they must reflect and actively engage with them. It is no longer sufficient to see Open Science as an external development that the universities merely monitor as a modern trend. Open Science is a reality and provides an advanced research framework, where

1 Text adopted from the Preamble of Open Science Strategy MU 2022–2028

2 For more details see e.g. https://council.science/current/news/unesco-science-commission-adopts-open-science-recommendation/

3 An open access to scientific publications in the electronic form without limitations on their use.

4 One of the visions in the Masaryk University Strategic Plan 2021–2028.

a research institution behaves by the Open Science principles. Open Science is an advanced environment for managing and disclosing the university's findings, an instrument for scientific diplomacy and a synergistic complement to communication, popularisation and transfer of knowledge and technologies. It is also important for social acceptance and perception of science and research as an integral part of societal responsibility and development.

As mentioned in the Strategic Research and Innovation Agenda (SRIA) of the European Open Science Cloud (EOSC): *"The digital age, the most recent stage in an evolving continuum of ways in which technology has supported science, presents an opportunity to improve the conduct of research in multiple directions, including with regard to openness, speed of access to scientific results, reproducibility and multi-disciplinarity. This should result in better science, increased trust in science, and an improved ability to meet global challenges. However, this potential will only be realised if research infrastructures evolve to allow scientists to exploit, in an easy-to-use and integrated environment, the (vast amounts of) relevant data being produced."* [5]

The focus should be on science as such and its excellence; Open Science and research infrastructures [6] have to be understood as indispensable support tools for its academic environment, and academic institutions should advocate that the Open Science principles to be a standard not only in the Czech Republic but also in the broader European and social context.

FROM OPEN ACCESS TO OPEN/FAIR DATA

The Berlin Declaration on Open Access to Knowledge in the Sciences and Humanitiesp [7] is currently the basic document for implementing Open Access. It builds on two previous initiatives: *Budapest Open Access Initiative* [8] and *Bethesda Statement on Open Access Publishing* [9]. These three documents are collectively referred to as the *Open Access BBB initiatives*.

The Berlin Declaration was adopted in October 2003 at a conference organised by the Max Planck Society in Berlin, was signed by 497 institutions worldwide, and has been widely adopted and matured to the de facto standard requirement of the European project calls [1].

However, the scientific publications alone are insufficient in data-driven 21st century science. There is an increased need to make available also the data that stay behind the publication – they are critical for reproducibility and thus trust in the published findings. But even that is not sufficient, as the data themselves become a valuable resource, opening opportunities for more research. Strong motivation for making research data available as a proof of the soundness of the published finding is the so-called *Replication crisis* [10]. Another is the principle of making the results of public money-funded research freely available. In

Europe, the Open/FAIR data requirements [3] started to emerge in the project calls [11] and currently are the standard requirement in projects [12].

OPEN/FAIR DATA, EOSC AND OPEN SCIENCE TODAY

Research data usually refers to the factual information or evidence generated during research (research project). Another definition emphasises the use, considering it as "data used for research" [13]. With the emphasis on data sharing and open access to the research data, the scientific publishers widely adopted the concept of data supporting a journal article. The *"data, including associated metadata, needed to validate the results presented in scientific publications"* are nowadays a standard component of journal articles, giving readers at least a theoretical opportunity to check the data used. However, this should be considered as the first step only, as this degrades the research data to a secondary position of "supporting the publication" only.

The interest is shifting towards **individual data sets** (*"other data, including associated metadata connected to a particular research project, as specified in the data management plan"*) [14], made available directly, without direct association to any specific publication. The data are becoming first-level citizen in the research space, having identity independent from any other scientific artifact. The datasets are assigned permanent identifiers and can be cited as any other scientific publication.

The **Open Research Data** are governed by the same principles as open data of the public sector explained in the previous chapter. We can consider the data supporting open-access articles as a classical representative of this concept data. However, the Open Research Data concept is too generic and cannot be applied to all research data (e.g. because they are protected by laws or represent valuable industrial property). Therefore, the FAIR data concept has emerged as a usable compromise between full openness and the restrictions that can accompany research data. In general, *"the FAIR Principles put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals"* [2,14].

By adhering to these features, FAIR data aims to enhance the value and impact of research data by enabling its widespread discovery, access, integration, and reuse, thereby fostering collaboration, reproducibility, and innovation in the scientific community.

European Open Science Cloud (EOSC) is an ambitious initiative by the European Union (EU) to create a digital platform that enables seamless access to research data, services, and infrastructure across Europe. The main goal of EOSC is to promote open and collaborative research, making scientific knowledge more accessible, transparent, and reproducible. The EOSC origins can be traced back to the realisation that scientific research has become increasingly data-intensive and requires advanced computational tools and infrastructure. The European Commission launched the European Cloud Initiative in 2016. It led to the establishment of EOSC as a long-term vision for the digital transformation of science and research in Europe, making it a unique horizontal partnership within the current HEU framework. By providing a federated and trusted environment, EOSC aims to enable researchers to access and share data, tools, and services across borders, disciplines, and sectors. The European Open Science Cloud (EOSC) and FAIR data share a close relationship, as both initiatives align their goals and principles regarding open and accessible data in the research community.

EOSC and FAIR data share a very close relationship, as both concepts are aligned in their goals and principles regarding open and accessible data availability in the research community. EOSC is a key instrument to deliver Open/FAIR Data in everyday practice in academia. It is supposed to "deliver Europe's contribution to the realisation of scientists', and science's, potential in the digital age, enhancing Europe's leadership position in exploiting digital capabilities at the service of science." 30 The three main objectives of EOSC are visualised in Figure 9. They are framing the current research data sharing landscape within Europe and together with the European Data Strategy and other "data spaces" they create the new generation of European digital infrastructure [5].

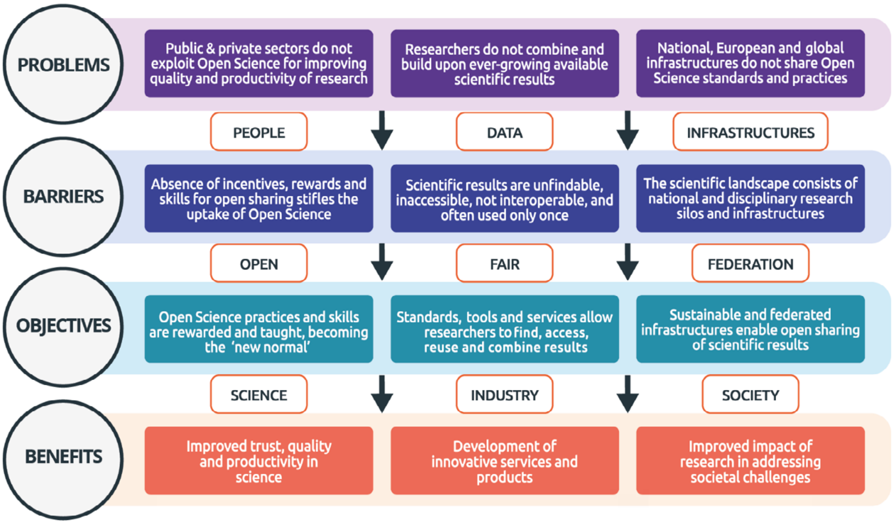

Figure 9: European Open Science Cloud Objectives Tree

The EOSC objectives show EOSC as an initiative that aims to federate and make available (although not necessarily fully open) research data across disciplines, sectors, and countries. To achieve this, EOSC promotes the adoption of open and FAIR data principles to enhance the accessibility and reusability of research data. While creating federated infrastructures for seamless access and collaboration, EOSC also seeks to support the development and provision of a wide range of services and tools for effective data management, analysis and sharing. It focuses on ensuring interoperability and seamless access to research data and services while respecting privacy, security, and ethical considerations. Last but not least, EOSC fosters collaboration and networking among researchers and institutions across Europe and beyond. EOSC can be considered the primary tool to foster an open science culture, supporting the transition towards openness and broad multidisciplinary research collaboration.

Research data are used for more than just research itself. The use of research data for initiatives like Citizen Science or Science Diplomacy is getting more and more attention, and the Policymakers are looking towards research and objective scientific information as one of the tools in the new post-factum era as was seen during the COVID-19 pandemic. The following chapters will dwell more in detail on concrete sectors of open science/research data management, that are key in today's world of data analytics, interactive visualisations, and web applications as a basic infrastructure to effectively support decision-making on various policy levels.

E-INFRASTRUCTURES AS A KEY STAKEHOLDER IN THE RESEARCH DIGITAL AGE

RESEARCH INFRASTRUCTURES IN THE EUROPEAN CONTEXT

Scientific communities thrive within a dynamic ecosystem centred around universities and research infrastructures. Universities provide an ideal setting for scientists to concentrate on their research and teaching. In contrast, research infrastructures establish networks of scientists within specific fields across countries, fostering cooperation and enabling domain-specific innovation. These infrastructures facilitate collaboration and distribution of services, irrespective of national or university affiliations.

International collaboration is one of the crucial aspects of research infrastructure operations. In Europe, the European Strategy Forum on Research

Infrastructures (ESFRI) identifies over 60 pan-European research infrastructures[5]. Additionally, European National Research and Education Networks (NRENs) have emerged as another supportive environment for research, education, and innovation, united under an association called GÉANT. These infrastructures, along with other institutions that bring together researchers from various scientific domains, continually develop and maintain robust environments for data storage, processing, and analysis—fundamental to modern research endeavours.

The European Open Science Cloud (EOSC) initiative has recently strengthened the cooperation among these infrastructures and institutions. EOSC aims to provide a federated and open multidisciplinary environment, enabling European researchers, innovators, companies, and citizens to publish, discover, and reuse data, tools, and services for research, innovation, and educational purposes. The data within EOSC adhere to the FAIR principles, ensuring they are Findable, Accessible, Interoperable, and Reproducible. Beyond data management, EOSC also encompasses data processing, as the value of scientific data lies in their manipulation, analysis, and application to real-world problems.

The history of the development of research infrastructures, with their robust technological background for storing and processing scientific data, together with new, data-centric initiatives such as EOSC, and challenges stemming from globalised society relying more on data collection and analysis, together represent a fertile environment for new, ground-breaking discoveries across scientific domains.

RELEVANT INFRASTRUCTURES FOR THE MEDICAL SECTOR

The above-mentioned European Infrastructures recognised by ESFRI can be categorised into six thematic areas[6], two of which are most relevant to this book.

First, **data, computing & digital research infrastructures**, commonly known as e-infrastructures. The only already established European e-Infrastructure recognised by the ESFRI is PRACE (Partnership for Advanced Computing in Europe), with three newly emerging infrastructures (EBRAINS, SLICES, SoBigData) in the preparation phase. However, the field of digital support of research concentrated in RIs is complemented by European Grid Infrastructure (EGI), European High Performance Computing Joint Undertaking (EuroHPC JU), EUDAT, OpenAIRE and GÉANT, which itself coordinates the operation of National Research and Education Networks (NERNs). All these infrastructures and collaborations seek to develop and maintain digital services in terms of

5 For more details see https://roadmap2021.esfri.eu
6 See ibid

networking, computing, and data management and foster the emergence of Open Science practices.

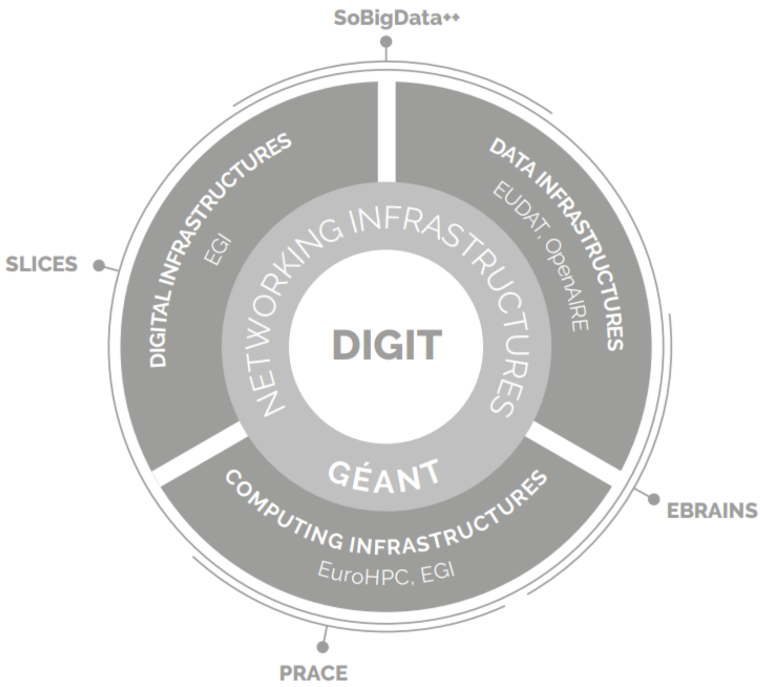

Figure 10: The landscape of the data, computing & digital research infrastructure domain

Second are infrastructures falling into groups labelled as **health & food**, specifically the health subdomain. Already existing infrastructures falling within this category are AnaEE, BBMRI ERIC, EATRIS ERIC, ECRIN ERIC, ELIXIR, EMBRC ERIC, ERINHA, EU-OPENSCREEN ERIC, Euro-BioImaging ERIC, INFRAFRONTIER, INSTRUCT ERIC and MIRRI, with four newly emerging RIs: EIRENE RI, EMPHASIS, IBISBA, METROFOOD-RI [15].

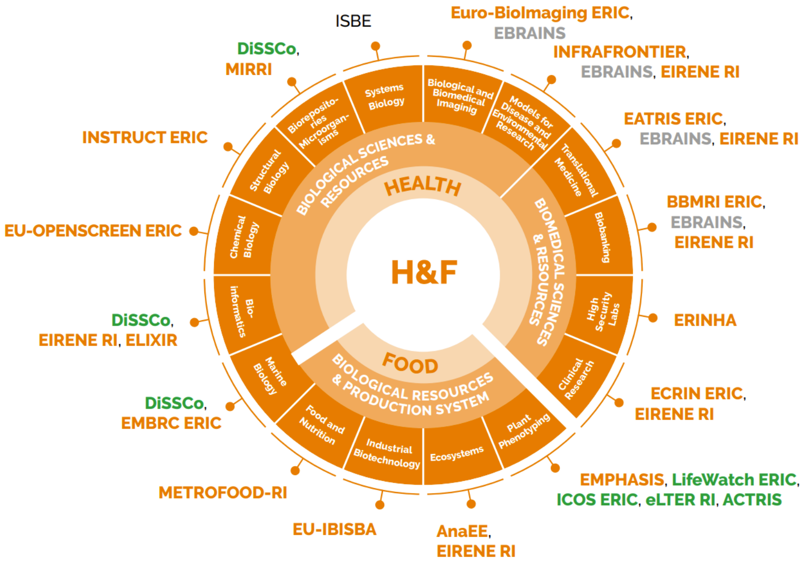

Figure 11: The landscape of the health & food domain

Within the health and food thematic area, the health subdomain encompasses many research areas, portfolio technologies, and related services. The dynamic nature of this domain is particularly noteworthy, especially considering the lessons learned during the SARS-CoV-2 pandemic. The response of research infrastructures (RIs) to the COVID-19 crisis is a remarkable demonstration of coordinated efforts in addressing an urgent problem.

The most prominent mission of RIs falling within this category is dealing with cancer and cardiovascular diseases, as they are responsible for an alarming number of total deaths in Europe. These RIs also seek to protect human health by increasing preparedness for and capability to respond to highly pathogenic infectious threats. Other strategic goals of these infrastructures are related to understanding plant performance in the context of health and famine challenges, increasing understanding of oceans, seas and inland waters, and many other related areas.

European research infrastructures continuously seek to support research via environments of services, tools and competencies required for domain-specific conduct of research. These requirements are increasingly more concentrated around data and various digital artefacts. With their inherent cross-sectional function, research infrastructures play a vital role in data management, from data storage, access, analyses, sharing and cooperation to data visualisation and publication for scientific communities and the general public.

SITUATION IN THE CZECH REPUBLIC

Following the clustering established at the European level, Roadmap of Large Research Infrastructures of the Czech Republic recognises six clusters of research infrastructures operated at and funded on the national level [16].

Ten research infrastructures fall within the health & food category on the national level, most of which also constitute a local node of European and pan-European infrastructures. These are the Bank of Clinical Specimens (BBMRI-CZ), the Czech Centre for Phenogenomics – (CO), the Czech Infrastructure for Integrative Structural Biology – (CIISB), the Czech National Node to the European Clinical Research Infrastructure Network – (CZECRIN), the National Research Infrastructure for Biological and Medical Imaging – (Czech-BioImaging), the National Infrastructure for Chemical Biology – (CZ-OPENSCREEN), the Czech National Node to the European Infrastructure for Translational Medicine – (EATRIS-CZ), the Czech National Infrastructure for Biological Data – (ELIXIR-CZ), the Infrastructure for Promoting Metrology in Food and Nutrition in the Czech Republic – (METROFOOD-CZ) and the National Center for Medical Genomic – (NCMG).

In contrast to the European landscape, where several e-infrastructures (e.g. GEANT, EGI.eu, EUDAT, PRACE, EuroHPC) co-exist, there is only one e-Infrastructure on the national level. e-INFRA CZ was created by merging three formerly independent infrastructures: CESNET (Czech NREN), CERIT-SC at Masaryk University, and supercomputing centre IT4Innovations at Technical University Ostrava. This merger provides a fertile ground for integrating the national infrastructure of resources and services, which has already proved to bring fruit in the Czech response to the European Open Science Cloud (EOSC) initiative on a national level. The Czech approach to the EOSC implementation is based on developing the National Data Infrastructure (NDI) with the National Repository Platform (NRP) in its heart. The e-INFRA CZ coordinates key research institutions across the national landscape to join forces in implementing EOSC and NDI.

The existence of a single national e-infrastructure guarantees an integration of the NDI and NRP into the already established environments focused on computing, networks, cybersecurity, and other topics related to IT in the research domain. Moreover, a single national e-infrastructure represents the best chance for the future continuation and development of the developed environment, services and practices, including those established within the Czech response to the EOSC initiative.

HOW E-INFRASTRUCTURES ARE CONNECTED TO THE DATA PROCESSING TOPIC

All of the latest progress in science and our understanding now relies on a steady increase of computational power dedicated to processing vast amounts of scientific data. Most recent and dramatic discoveries in various science fields were made due to advancements in the computing area. Optimisations in numerical methods and the remarkable growth of available computational power reflect many exceptional advances in research, such as climate modelling, protein folding, drug discovery, and energy research. It is fair to say that furthering our (human) theoretical understanding has transformed from "pen-and-paper" to designing simulations and modelling applying mathematical concepts using technologies offered by accomplishments within the computer science area.

Research infrastructures at both national and international levels aim to offer baseline data-related services and appliances to the scientific community. This is especially true for e-infrastructures, which, in contrast with domain-specific infrastructures, are focused on administering the raw power and generally applicable IT environments. Research infrastructures represent an indispensable conduit for modern research and play a crucial role in the national and international ecosystem related to the management and processing of research data.

DATA MANAGEMENT PLANNING AT RESEARCH INSTITUTIONS

DATA MANAGEMENT PLANNING AS A SKILL FOR EVERYDAY RESEARCH, TECHNICIAN AND STUDENT

Having standard requirements on projects to produce FAIR data, the researchers and project-support teams in their institutions are naturally motivated to consider their implementation. At the same time, there is a self-evident lack of these competencies across the academic institutions in the Czech Republic and the general research community.

In reaction to the lack of skills on the one hand and requirements from project funders on the other, Open Science support teams started to emerge in large universities in the Czech Republic[7]. At the centre of their focus lies aggregating

7 Masaryk University (https://openscience.muni.cz/en), Charles University (https://openscience.cuni.cz/OSCIEN-1.html)

knowledge regarding project call requirements, best practices in data management, repositories available for various scientific disciplines for publishing data, and, very importantly, the transfer of relevant information and knowledge to the research community.

Their members started self-organising in so-called expert groups to improve skills and share knowledge among the Open-Science-support teams. Just to name the most prominent of these initiatives, informal Discord groups *AKVŠ-PS Open Science, datasteward.cz*, and *Data Stewardship Wizard* already successfully interconnected many data experts across the Czech Republic.

TOOLS FOR DATA MANAGEMENT PLANNING

One of the first concrete requirements encountered by the researchers while interacting with the research funding organisations is creating and keeping an up-to-date Data Management Plan (DMP). This requirement logically follows the intention of research funding organisations (RFOs) on handling research data as first-class-citizen research outputs: DMP should certify that the project team has a clear idea of how to handle the data from creating to processing, preserving, and finally sharing.

To support DMP preparations, a variety of tools was created. DMPonline[8] represents a tool following document templates from RFOs but transforming them to electronic form with guidance and helping texts from Open-Science-support teams of various institutions. More advanced tools were developed later. For example, ARGOS[9] follows a similar concept as DMPonline in transforming DMP templates into electronic forms in specialised information systems. However, it divides the description of a research project from the description of a dataset to allow common practices of using one dataset in multiple projects and multiple datasets in one project. ARGOS also benefits from integration with other relevant systems like OpenAIRE Research Graph[10] or Zenodo[11] to transfer data in standard forms. The possibility of systems and data integration is one of the motivation factors for using specialised information systems for DMP support instead of creating and sharing them as standard documents in common office formats.

The most advanced DMP-supporting tool, from my point of view, is Data Stewardship Wizard (DSW)[12]. In contrast to the previous tools, its primary intention is to support data management *planning* as a process, not only creating data management plans as documents. DSW builds on an expert system, defining

8 https://dmponline.dcc.ac.uk/
9 https://argos.openaire.eu/
10 https://graph.openaire.eu/
11 https://www.zenodo.org/
12 https://ds-wizard.org/

the so-called knowledge model (KM) for a particular scientific discipline. The KM should describe data handling in the area, regardless of the specific DMP template. DMP documents are generated according to a particular DMP template on request for the information provided according to KM.

DSW is focused on modern implementation, supporting real-time collaborations of multiple persons and integration with other information systems providing rich API covering the tool's full functionality. The principle of the wizard highly improves the user experience: The user is asked only the necessary questions. More details are required by the system only if relevant to the current answers. Where suitable, the system is connected to vocabularies and ontologies to support machine-actionability on the DMP data. Metrics of fulfilling FAIR and best-practice requirements are defined in the system and indicated for every DMP to provide users feedback on the quality of the data management in the described project.

LESSONS LEARNED FROM DMP AT CZECH ACADEMIC INSTITUTIONS

Recent experience from Czech academic institutions shows a clear need for training research-data-management support staff like data stewards. There is a lack of mature description of the required competencies. Still, experience from the Open Science support departments of the Czech academic institutions shows the need for knowledge in data management planning, data handling and their specifics in various disciplines, specific IT skills (handling of large datasets and sensitive data, real-time data processing, support of AI computations and accelerations, design of modern applications using containerisation, ...) and expertise in data storage, discipline-specific repositories and metadata creation and management.

The first contact with the Open Science support team is usually initiated by the researchers not knowing how to handle FAIR data / DMP requirements in the project they are participating in or not knowing how to handle sensitive data in their project following the current best practice and project and law requirements.

Handling of sensitive data usually requires the use of dedicated infrastructure. Some of them are available or under construction in the Czech Republic (for example, CERIT-SC SensitiveCloud[13]). Handling DMP is usually easier – Open Science support teams can provide consultation by filling in the RFO DMP template and providing support tools like DSW instance.

13 https://www.cerit-sc.cz/infrastructure-services/sensitivecloud

For example, at Masaryk University, we have our own instance of DSW, called DSW MUNI[14], that is integrated with an institutional information system that, besides other responsibilities, is used to manage research projects. This integration, made possible via full-featured DSW API, allows MU researchers and management to keep precise track of data management of every solved project by binding DSW DMPs with current research projects. On the other hand, our experience also shows that the researchers often make DMP preparations shortly before the deadline for their release for RFO. And in this case, they sometimes prefer filling in the RFO DMP template directly than filling in the full description of data management in their research in DSW. That indicates that the researchers are still thinking about DMPs as 'bureaucratic' documents for RFO that the process that can help them improve, especially in long-term, overall data handling in their projects.

Another problem often presented by the researchers is the lack of suitable data repositories to store research data for fulfilling the FAIR data requirements of RFO. Researchers from MU CEITEC MAFIL[15], specialising in human magnetic resonance imaging, are currently implementing support for producing FAIR data as the standard output of their laboratory for all their measurements. The investigation of suitable repositories for publishing their data indicates OpenNeuro as the best appropriate existing repository. However, it still does not fit their needs fully.

For example, experts from the Masaryk University CSIRT team[16] were even less successful when looking for a suitable repository for publishing their dataset Encrypted Web Traffic: Event Logs and Packet Traces [17]. As there was a lack of an appropriate topic repository, they tried to use Zenodo. However, the size of the dataset – 270 GiB – was far beyond the available limits of Zenodo. We decided to take advantage of Masaryk University's ability to assign Digital Object Identifier (DOI) [18] persistent identifiers for datasets on our own: The dataset was assigned DOI and stored in the university's cloud storage, making DOI resolve to read-only share link. This is considered a short-term solution; in the long term, we expect to move the dataset to a national data repository (a pilot version of the Czech data repository is currently under development[17]) or use a newly emerged suitable topic repository. As DOI is a persistent identifier, the users will not be affected by the change of the dataset's location as we will guarantee the DOI will always resolve to the current dataset's location. This demonstrates the importance of one of the FAIR data requirements – persistent identifiers hiding users from the complexity of physical data storage and providing them with the

14 https://dsw.muni.cz/
15 https://mafil.ceitec.cz/en/
16 https://csirt.muni.cz/?lang=en
17 https://data.narodni-repozitar.cz/.

functionality identifier of the dataset itself, guaranteeing access to the current dataset location.

Yet another possibility how to solve the lack of a suitable data repository is to build your topical repository. Even though the best practice is to use an established repository, it is still a valid approach if none exists, for example, in the case of emerging new disciplines. We also have experience with this approach: Our colleagues from the Faculty of Science of Masaryk University operate the World Spider Trait database[18]. To enhance the FAIRness of the contents of this repository, their operators integrated with our institutional DataCite account for automated assigning of DOIs for records in this database[19].

REFERENCES

[1] European Commission. Open Science [Internet]. European Commission; 2020 [cited 16 Jul 2023]. Available from: https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science_en.

[2] Wilkinson M, Dumontier M, Aalbersberg I, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016;3:160018.

[3] GO FAIR Initiative. FAIR Principles [Internet]. GO FAIR Initiative; 2023 [cited 21 Jul 2023]. Available from: https://www.go-fair.org/fair-principles/.

[4] European Commission, Directorate-General for Research and Innovation. Towards a Reform of the Research Assessment System: Scoping Report [Internet]. Luxembourg: Publications Office of the European Union; 2021. Available from: https://data.europa.eu/doi/10.2777/707440.

[5] European Open Science Cloud. Strategic Research and Innovation Agenda (SRIA) of the European Open Science Cloud [Internet]. 2022 [cited 21 Jul 2023]. Available from: https://eosc.eu/sites/default/files/SRIA%201.1%20final.pdf.

[6] Ministry of Education and Research of Sweden. Lund Declaration on Maximising the Benefits of Research Data [Internet]. 20 Jun 2023 [cited 21 Jul 2023]. Available from: https://www.government.se/contentassets/69ab5c1102d5435ea24ce18da837b17d/lund-declaration-on-maximising-the-benefits--of-research-data.pdf.

18 https://spidertraits.sci.muni.cz/

19 https://commons.datacite.org/repositories/eqjcw9y

[7] Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. Open Access Initiatives of the Max Planck Society [Internet]. Berlin: Max-Planck-Gesellschaft; 22 Oct 2003 [cited 21 Jul 2023]. Available from: https://openaccess.mpg.de/Berlin-Declaration.

[8] Budapest Open Access Initiative: The original Declaration and guidelines to make research free and available to anyone with internet access and promote advances in the sciences, medicine, and health [Internet]. Budapest: Budapest Open Access Initiative; 2023 [cited 21 Jul 2023]. Available from: https://www.budapestopenaccessinitiative.org/.

[9] Bethesda Statement on Open Access Publishing. Digital Access to Scholarship at Harvard [Internet]. Harvard: Harvard University; 20 Jun 2003 [cited 21 Jul 2023]. Available from: https://dash.harvard.edu/bitstream/handle/1/4725199/Suber_bethesda.htm.

[10] Baker M. 1,500 scientists lift the lid on reproducibility. Nature. 2016;533:452–4.

[11] Open Research Data (ORD) – the uptake in Horizon 2020 [Internet]. European Commission, Directorate-General for Research and Innovation; 2018 [cited 21 Jul 2023]. Available from: http://data.europa.eu/88u/dataset/open-research-data-the-uptake-of-the-pilot-in-the-first-calls-of-horizon-2020.

[12] Horizon Europe (HORIZON): Programme Guide [Internet]. Version 3.0. Europe: European Commission; 1 April 2023 [cited 16 Jun 2023]. Available from: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/programme-guide_horizon_en.pdf.

[13] Koščík M, Polčák R, Myška M, Harašta J. Research data and their legal context. In: Koščík M, Polčák R, Myška M, Harašta J (eds). Research Data and Research Databases. Legal Framework for the Processing and Sharing of Scientific Findings. Prague: Wolters Kluwer; 2018. p. 17–24.

[14] H2020 Programme AGA – Annotated Model Grant Agreement. Version 5.2 [Internet]. European Commission; 26 June 2019 [cited 14 Jul 2023]. Available from: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/amga/h2020-amga_en.pdf.

[15] ESFRI Roadmap 2021 – Strategy Report on Research Infrastructures [Internet]. European Strategy Forum on Research Infrastructures; 2021 [cited 14 Jul 2023]. Available from: https://roadmap2021.esfri.eu/media/1295/esfri--roadmap-2021.pdf.

[16] Roadmap of Large Research Infrastructures of the Czech Republic for the years 2016–2022 [Internet]. Ministry of Education, Youth and Sports of the Czech Republic; 2019 [cited 14 Jul 2023]. Available from: https://www.vyzkumne-infrastruktury.cz/wp-content/uploads/2019/11/Aktualizace-Cestovn%C3%AD-mapy-2019_en.pdf.

[17] Špaček S, Velan P, Čeleda P, Tovarňák D. Encrypted Web traffic dataset: Event logs and packet traces. Data Brief. 2022;42:108188.

[18] What is a DOI. DOI Foundation [Internet]. DOI Foundation; 2022 [cited 1 Jul 2023]. Available from: https://www.doi.org/the-identifier/what-is-a-doi/.