Disclaimer: **This deliverable has not yet been reviewed by the European Commission. Its content might therefore change as a result of the view process.**

D6.2: Final analysis and mapping of existing European and national OS infrastructures with regard to promoting responsible OS

Authors: Carole Chapin, Nathalie Voarino

Editor: Olivier Le Gall

Responsible Open Science in Europe

ROSiE

Grant Agreement no.: 101006430
Lead contractor for this deliverable: Ofis-Hcéres

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Deliverable factsheet:

Project Number:	101006430
Project Acronym:	ROSiE
Project Title:	Responsible Open Science in Europe
Title of Deliverable:	Final analysis and mapping of existing European and national OS infrastructures with regard to promoting responsible OS
Work Package:	WP 6
Due date according to contract:	31.01.2023
Editor(s):	Olivier Le Gall
Contributor(s):	Carole Chapin, Olivier Le Gall, Nathalie Voarino
Reviewer(s):	François Jost, Tom Lindemann, Elina Koivisto & Rosemarie de La Cruz Bernabe
Approved by	Rosemarie de La Cruz Bernabe

ABSTRACT:	This report gives an overview from a mapping exercise to collect information on the existing OS platforms and analyse their needs and existing weaknesses, with respect to actively pursuing open approaches in science and research, while complying with legal frameworks and ethical standards.
Keyword List:	Research Integrity, Open Science, Open Access, Research infrastructures, EU, FAIR Principles

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Consortium:

	ROLE	NAME	Short Name	Country
1.	Coordinator	UNIVERSITET I OSLO	UiO	Norway
2.	Partner	ÖSTERREICHISCHE AGENTUR FÜR WISSENSCHAFTLICHE INTEGRITÄT	OeAWI	Austria
3.	Partner	VEREIN DER EUROPÄISCHEN BÜRGERWISSENSCHAFTEN	ECSA	Germany
4.	Partner	EUREC OFFICE GUG	EUREC	Germany
5.	Partner	TIETEELLISTEN SEURAIN VALTUUSKUNTA	TSV	Finland
6.	Partner	HAUT CONSEIL DE L'EVALUATION DE LA RECHERCHE ET DE L'ENSEIGNEMENT SUPERIEUR	Hcéres	France
7.	Partner	INSTITUT NATIONAL DE RECHERCHE POUR L'AGRICULTURE, L'ALIMENTATION ET L'ENVIRONMENT	INRAE	France
8.	Partner	NATIONAL TECHNICAL UNIVERSITY OF ATHENS	NTUA	Greece
9.	Partner	UNIVERSIDADE CATOLICA PORTUGUESA	UCP	Portugal
10.	Partner	LATVIJAS UNIVERSITATE	UL	Latvia
11.	Partner	TARTU ULIKOOL	UT	Estonia
12.	Partner	UNIVERSITETET I SOROST-NORGE	USN	Norway

Revision history:

VERSION	DATE	Revised by	Reason
0.1	23.12.2022	Carole Chapin, Olivier Le Gall	First draft of workshop results
0.2	31.01.2023	Olivier Le Gall, Nathalie Voarino	First draft of entire report
1.0	01.02.2023	Consortium	Final draft
2.0	17.02.2023	Rosemarie de La Cruz Bernabe	Final approved report

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Table of contents

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Analysis and mapping of existing European and national Open Science infrastructures

Challenges and solution for responsible Open Science in Europe

1 Introduction

One of ROSiE's objectives is to provide all open science (OS) actors (researchers, existing and future platforms, decision makers, etc.) with customized solutions proposed through a knowledge hub. The knowledge hub, designed and governed to ensure its sustainability beyond termination of ROSiE, will support all research disciplines in providing appropriate knowledge to support open approaches in science and research while complying with legal frameworks and ethical standards. In the ROSiE context, there is also a particular interest to the readiness of existing OS platforms to handle sensitive personal data, as well as the potentialities of various technologies including several versions of distributed ledger technologies (the blockchain). ROSiE's knowledge hub will be designed and produced after assessing which of the existing technologies (including blockchain, permissioned or permissionless, public or private) are the most appropriate to share and protect personal data and, more generally, to promote RE/RI, considering the particularities of existing OS platforms (e.g. disciplinary focus, level of citizen science involvement, prevalent types of data, number of involved users, among others). Different candidate technologies may be developed into appropriate infrastructures to ensure responsible OS, including GDPR compliance, accountability, reliability and consent, data storage and security measures, etc.

For this purpose, we examined how the OS infrastructures should be designed, how to organize and manage different solutions, and how knowledge on RE/RI and OS can best be delivered to the OS infrastructure executives and other stakeholders through the ROSiE Knowledge Hub.

This report gives an overview from a mapping exercise to collect information on the existing OS platforms and analyse their needs and existing weaknesses, with respect to actively pursuing open approaches in science and research, while complying with legal frameworks and ethical standards. This analysis has been conducted in collaboration with partners from other ROSiE WPs, especially WP1 and WP2 for mapping and analysis of ethical, research integrity, social and legal dimensions, and WP3 and WP4 for stakeholder and community consultation respectively. Two series of workshops enabled consultations of experts from OS infrastructures covering a diversity of thematic, disciplinary, geographical and other contextual origins.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

1.1 The main current and forthcoming ethical, research integrity, social and legal concerns

The mapping analysis relies on the findings presented by the reports conducted by ROSiE to identify the main current and forthcoming ethical, research integrity, social and legal concerns about implementing OS strategies. This section summarizes the results from these reports. **In the following, findings from the cited reports will be outlined mostly verbatim and partly paraphrased, with minor additions where seen appropriate.**

The ROSiE "Report on the relationship (tensions, challenges, overlaps) between RI, the wider RE perspective and OS" (August 2022[1]) analyses how the different official codes of conduct on research integrity across Europe approach open science objectives. The goal was to analyse how the official RI codes in different European countries (EU+) approach OS objectives and whether RI/RE and OS principles match and to what extent in these codes. It underlines that responsibility in research is spread across many areas (research ethics, research integrity, open science, responsible research and innovation, science communication) and that these areas are usually treated independently.

These findings raise questions about how the different areas are operationally understood and implemented in the research infrastructures funded by the European Union or member states. In the construction of research infrastructures and related tools, is there the same separation? Can this be related to the way in which codes and guides are presented, which therefore do not provide the appropriate framework and incentive?

The ROSiE "Report on social challenges and implications related to Open Science" (July 2022[2]) explores different kinds of challenges. Especially, it underlines the social challenges brought by OS processes, the attitudes towards OS in the scientific communities and the roles of different actors in the process of implementation of requirements of research ethics and research integrity in the context of OS.

One of significant social challenges for integration of RI / RE and OS into national research and innovation schemes emphasized by several project reports is **economic disparities between countries** (both between EU countries and globally). The economic disparities between countries have direct consequences on their implementation of OS. On the one hand, OS may bring economic benefits for countries, societies, and institutions, while on the other hand, several reports mentioned costs emerging in the context of OS, which can be a significant economic barrier for developing countries and institutions experiencing financial struggles.

Cultural differences in the context of scientific practice are also a major challenge, not only between different countries but also between different groups in the same country. The degree of

[1] Carvalho, A.S. & Strecht Almeida, M. (2022): D.1.1: *Report on the relationship (tensions, challenges, overlaps) between RI, the wider RE perspective and OS*. ROSiE: Responsible Open Science in Europe, unpublished project deliverable.

[2] Mežinska, S.; Mileiko, I.; Neiders, I. & Kalēja, J. (2022): D.2: *Report on social challenges and implications related to Open Science*. ROSiE: Responsible Open Science in Europe, unpublished project deliverable.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

awareness of OS and attitudes towards it varies across scientific disciplines. When looking at cultural barriers for implementation of OS, academic culture seems to be one of significant barriers for research funding organisations (RFOs) in their efforts to introduce OS. Some studies show that there can be a cultural resistance towards re-organization of scientific process, also specifically in case of OS significantly changing the existing ways of doing research[3]. In case of cultural resistance, a targeted management to change academic culture at institutional level, as well as personal attitudes becomes a crucial incentive to promote OS. Cultural issues also include a tension between values of openness and transparency vs. intellectual property and sensitivity of research data.

Another challenge is **the existence of a gap between society and science and the need to increase public trust in science** and its products, including shared knowledge**.** OS has been suggested as one of the approaches implying narrowing this gap and building trust by increasing openness to external actors. The public engagement in science is fostered to bridge the gap between researchers and the public, to integrate public concerns in research practice and to disseminate research results. At the same time, the existing diversity of approaches and definitions for public engagement may cause confusion in scientific communities. OS and open access publications are considered as a part of the solution in this context. However, there still are a number of difficulties even in countries where the OS principles have been included in policy documents: publications are written in specific jargon and may not be understood by the general public – not to mention non-native English-speaking public-; it is difficult to change existing publishing practices, etc. Even if data and publications are technically available to public, there may be barriers to the re-use or to free use of knowledge.

High competition for research funding, existing inequalities in scientific communities and other factors leading to **hyper-competitive practices in the field of science** increase potential of breaches of research integrity. Therefore, there is an identified "hidden opposition between excellence and RRI, viewed as two competing priorities[4]" and the fact that the existing system for assessing the scientific performance promotes publishing in English language journals with high impact factor and fast publication schedule, rather than prioritizing social impact, openness and responsibility towards the society. Accordingly, in such an environment open access may be viewed as an additional challenge for achieving excellence in science by researchers. Furthermore, the implementation of OS requires a lot of cooperation among scientists and institutions, and it is hardly possible to introduce OS in a hyper-competitive environment.

There are differences between countries in scientists' attitudes towards OS, possibly caused by various socioeconomic and/or cultural factors. There are also differences in **attitudes and readiness to engage in OS between different fields of science**. Differences between scientific disciplines are explained by differences in technical skills, traditions, data specificity (qualitative/quantitative/sensitive/personal etc.), history of practicing OS, etc. Ethical issues in case of human subject research may affect researchers' ability (and sometimes willingness, according to the ROSiE report) to share data. Medical and social science researchers involving human participants

[3] Martin, E. G., & Begany, G. M. (2017). Opening government health data to the public: benefits, challenges, and lessons learned from early innovators. Journal of the American Medical Informatics Association, 24(2), 345-351. Cited by Mežinska, S.; Mileiko, I.; Neiders, I. & Kalēja, J. (2022), in ROSiE's report, *op.cit*.
See also: Maire, A. The digital disruption of science: Governments and scientists toward an "Open Science". The Digital Transformation of Labor, 2019. Armeni, K et al., "Towards wide-scale adoption of open science practices. The role of open science communities, 2021.

[4] d'Andrea, L., Berliri, M. and Federico M. (2018). Summary Report, Deliverable 1.3. Available: https://zenodo.org/record/1434355, p. 82. Quoted by ROSiE D.2 report, *op.cit.*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

sometimes cannot share research data because of confidentiality or privacy issues. Some scientific disciplines such as genetic genealogy, atmospheric science, and oceanography have well-developed traditions in OS and data sharing and have developed the necessary infrastructure and databases, whereas other disciplines may lack this experience, traditions and infrastructure. Publishing of preprints has a long tradition in physics, mathematics, astronomy, and information technology, nevertheless this practice is new for many other disciplines; the coming years will tell whether such a tradition spreads in biomedical research where the use of preprints emerged during the Covid-19 crisis. There might be differences in attitudes related to perception of data, where some scientists may almost see the research material and data as their personal property.

Different conceptualization of **responsibility** and various understandings of what responsibility means in the context of RRI and OS show that there are differences between countries, institutions and fields of science. Responsibility implies **close collaboration of all social actors and institutions involved in OS and RRI.** Responsible research at all stages of the Research & Development process involves individual responsibility of the researcher. Another group of responsible actors is **policy-makers and RFOs. Academic and research institutions** are important stakeholders with their responsibilities regarding responsible research and implementation of OS. **Academic publishers** are responsible for enforcing OS principles, as well as strengthening research ethics and integrity. Another important stakeholder in OS process is **industry**.

The ROSiE's reports[5] "Report on a strategy to engage Stakeholders" and "Report on [stakeholders'] interviews" outline the stakeholder engagement strategy of the ROSiE project. It contextualises stakeholder engagement within the overall work of ROSiE, expounds the aims of stakeholder engagement, specifies the types of stakeholder engagement formats, and provides an overview of the stakeholders identified in the stakeholder mapping. A major aim of stakeholder engagement is to ensure the inclusion of a diversity of perspectives in the exploration phase of the project. The results of this broad multi-perspective analysis informed the development of the mapping and analysis methodology. Before targeting specific groups, it was important to take into account the categories of stakeholders from whom insights were particularly expected. The list of stakeholders included:

- **Researchers,**
- **RPOs,**
- **Research Ethics Committees (RECs) and RI Offices (RIOs),**
- **RFOs and scientific journals,**
- **Research managers, research policymakers and advisory bodies,**
- **Science educators and journalists,**
- **Industry associations,**
- **Citizen science associations,**
- **Civil society organisations,**
- **The general public**.

[5] Lindemann, T. (2021): D.3.1: *Report on a strategy to engage Stakeholders*. ROSiE: Responsible Open Science in Europe. Lindemann, T.; Häberlein, L. & Hövel, P. (2022): Report on interviews. ROSiE: Responsible Open Science in Europe.
Available : https://rosie-project.eu/deliverables/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

It was also underlined that involving researchers from a broad range of countries in stakeholder engagement is important because most research infrastructures are funded, managed, and operated at the national or federal level, often embedded in national research strategies. Furthermore, it seems likely that **research infrastructures vary not only between countries but also between scientific disciplines**. Researchers embedded in research infrastructures that do not reward OS might be less inclined to engage with ROSiE than researchers embedded in research infrastructures that already incentivise OS. Research managers often serve as intermediaries (by being involved in grant applications and grant management, supporting researchers in meeting ethical and legal requirements, and assisting RECs and RIOs in their daily work) between the upper echelons of organisational governance in RPOs and researchers. **Especially ERI managers are well positioned to provide insights about the challenges of current OS practices** and can help in assessing whether proposed guidance materials are practically useful.

Interviews with stakeholders described how they conceptualise OS and which ethical, legal and policy issues they consider relevant.

Based on the results from previous interviews, one can expect that most participants targeted for the mapping and analysis workshops will view OS favourably and share many or all the values underpinning OS, such as availability and transparency of the research outcomes. At the same time, they identify significant challenges created by opening of science. OS means open access to knowledge for everybody, although it should not necessarily mean access without any restrictions **in case restrictions are justified and access mechanisms transparently described.** Training is crucial to support the transition to OS. Data curation is costly. There are also concerns about intellectual property rights and patents. Another challenge mentioned comes from **the arts and humanities**, where many concepts of OS seem not to be easily transferable, according to some stakeholders: for example, reproducibility in primarily interpretive methodologies. It was also pointed out that OS also creates new and exacerbates **existing research ethical challenges, especially in the realms of data protection, intellectual property rights and societal engagement with research**. All these challenges call for finding the right balance between promoting trust in research through openness and transparency on the one hand and safeguard privacy rights as well as legitimate interests of innovators on the other hand. The ethics of privacy protection in research are inherently intertwined with the legislation (GDPR, intellectual property law and patent requirements).

The need to create proper infrastructures for data management was underlined; however, technical aspects of infrastructure development were not named as a major concern because **existing infrastructural developments are perceived to be on the right track,** and technological progress and investments in platforms such as the European Open Science Cloud (EOSC) will further decrease technological barriers. According to the reports, training in responsible OS is desirable, and could be integrated in trainings in responsible research and good scientific practice, preferably **hands-on and practice-oriented rather than theoretical and general trainings**. Consequently, in the OS transition, research environment and data practices and management are closely related and cannot easily be analysed separately. Essentially **all aspects of OS related to open data are inherently linked to data practices.**

The move to OS also has created new challenges when it comes to **publishing and disseminating research**, albeit seemingly with some noticeable differences between different disciplines. This challenge might be particularly acute in disciplines where books are a major type of publication, and related to the problematic effects creative commons licences, such as CC-BY, can have : those can be republished in inadequate formats without the consent of authors or original publishers, as long as the text corpus remains unchanged. **Data stewards** on the institute or faculty level could be effective advisers because of their familiarity with disciplinary cultures and challenges. More generally, the reports cite various interviewees who alluded **to the importance of offering guidance on the**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

appropriate level within RPOs, which can be, depending on the case, research management, or operational actors at the institute or faculty level.

An important point of concern was raised, to the fact that efforts to promote responsible OS are somewhat hampered by the fact that **the RE/RI and the OS communities are currently separated**. Creating and strengthening bonding between these communities could help to increase synergies between RE, RI and OS.

When asked about how the ROSiE Knowledge Hub should ideally be designed to yield significant added value, many stakeholders specifically emphasised **the importance to ensure sustainability after the end of the project**: through updating mechanism or materials that have a high chance of long-term relevance. There is a risk that a new platform fails to create unique added value, therefore, as a new platform the ROSiE Knowledge Hub should have novel features and could also be integrated into existing platforms and linked to existing tools that support responsible OS, such as **tools to create proper data management plans or data anonymization tools.** The knowledge hub should be **user-friendly** and include **interactive elements**, such as a helpdesk function (creating some degree of collaboration while providing insights about stakeholder needs) or a decision-tree that incorporates different disciplinary perspectives.

As this summary illustrates, the findings presented in other ROSiE works provide good preparatory ground for the mapping and analysis task. They allow to identify priority angles (which actors and which issues) and to converge on useful learnings of the needs and challenges for responsible OS practice, in particular in view of the adaptation of the Knowledge Hub to these needs. The methodology adopted therefore takes into account these previous results.

1.2 Task 6.1. – Map and analyse the existing OS platforms and infrastructures

Existing OS platforms/infrastructures (e.g., those constructed under the aegis of and/or coordinated by the European Open Science Cloud (EOSC) or the national OS dynamics in the member states), have been examined in the light of the above concerns with regards to their implementation, to their engagement practices and to the SwafS community of practices. Their strengths, weaknesses, challenges and needs have been mapped and analysed, in relation with all aspects of responsible science and according to their relative backgrounds (types of services, thematic, academic disciplines, etc.).

It was anticipated that their overall level of responsibility regarding the OS platforms implementation would be high, given the following:

- the standards already established by the EC while implementing EOSC and while supporting infrastructures altogether, and similarly by the individual member states for their national infrastructures;
- the expertise and skills enrolled by all infrastructures during their respective set-up and running phases.

Therefore, the mapping and analysis enable two separate types of outcomes:

- Sharing strengths between infrastructures,
- Implementing innovative functions to promote responsible research.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

This listing, analysis and mapping was done together with stakeholders: at the European level with the governance of EOSC and specific OS infrastructures including those for citizen science.

The mapping and general analysis identified the strengths, weaknesses and needs of each OS infrastructure for the implementation of responsible OS, according to their respective contexts. Special attention was given, through a model in three categories of research practices (see below) to their disciplinary field and/or thematic coverage, the specific concerns attached (protection of personal data or intellectual property, reproducibility of data production/use/reuse, etc.).

The analysis has been conducted twice: first based on a first meeting of representative OS platforms, to provide a preliminary report, and for a second time, once the outputs of WP1, WP2 and WP3 were available, and a series of workshop was arranged to provide the final analysis and mapping of existing European and national OS infrastructures with regard to promoting responsible OS.

1.3 Inputs from the preliminary report

The final report is based on the results of the preliminary report[6]. Science relies largely and increasingly, especially for its opening process, on research infrastructures which are either (i) generic in their scope but dedicated to OS such as OpenAIRE, D4Science, etc., or (ii) thematically focused but with a commitment to opening scientific processes and especially the release of data and publications such as DARIAH-EU, BBMRI-ERIC and other ERICs (European Research Infrastructure Consortia).

During its first 18 months, ROSiE has mapped and analysed existing research infrastructures and their relation to, or their role in, promoting responsible OS. Most EC-supported research infrastructures have a solid commitment and experience in addressing responsible OS issues. ROSiE sees the analysis of these infrastructures' current practices as a robust first step towards its goal. For this purpose, an online workshop was organized in July 2021. In this first workshop, it was understood that "open science" is not exactly a product or an object, but a process: the opening of the entire scientific activity. This systemic transition takes some effort, and the European research community is committed towards the creation of the European Open Science Cloud (EOSC) "*to provide European researchers, innovators, companies and citizens with a federated and open multi-disciplinary environment where they can publish, find and re-use data, tools and services for research, innovation and educational purposes*[7]". While the first workshop was advertised during the EOSC symposium (June 2021), it was the organizers' choice to target a specific audience for this first workshop so that the focus was on research infrastructures and more specifically ERICs. The online workshop gathered a diverse panel of ERICs: of the twenty-two currently established ERICs, seven were represented, covering a diversity of scientific fields such as humanities, life sciences, environment sciences, etc. The ambition of this workshop was to identify the main describers of the situation, to be explored through additional focused workshops.

[6] Le Gall, O. & Medves, M. (2021): D 6.1: *Preliminary analysis and mapping of existing European and national Open Science infrastructures with regard to promoting responsible Open Science*: Responsible Open Science in Europe, unpublished project deliverable.
[7] quoting the EOSC Portal https://eosc-portal.eu/.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

During the preliminary workshop, most of the attention focused on two matters, (1) compliance with the GDPR (General Data Protection Regulation, EU 2016/679) and (2) informed consent.

Some specific questions required further examination, such as for instance “As an end-(re)user, how can I trust that the online data set which I am about to re-use is not the result of falsification or fabrication? How can I trust that it has not been produced against ethical rules such as informed consent of questionable validity or breaches to animal welfare, among others?”

It was also observed that much of the focus of the interviewees was on data. Data is only one of the products of science, although an illustrative one and one that gathers much of the concerns of end-users. However, it was recognized that the focus of the ensuing workshops therefore had to be enlarged to diverse aspects of responsibility in the context of creating and using different types of products of OS such as data, open access research tools, and resources.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

2 Methods

In research infrastructures in Europe, different levels of commitment, knowledge and tools on responsible OS can be observed. Especially, one can observe that there is not always a bridge between OS tools and processes directly labelled as research integrity or research ethics practices such as training, ethics committees, inclusiveness policies, etc. Moreover, experts on these topics and actions tend to be separated. Therefore, there is a need for documentation or information platforms and a better knowledge on responsible OS. This is the main challenge on conciliating RI and OS. Informative reports from different ROSiE WPs, with recommendations towards the Knowledge Hub, and the Knowledge Hub itself are the first steps to address these challenges.

Based on the list of relevant stakeholders, research infrastructures emerged as major interlocutors for this task. A list of infrastructures was drawn up, considering those specifically engaged in OS and those with active OS policies.

The two series of workshops included contributions from[8]:

- Directs contributions from all ROSiE WPs,
- OS platforms: EOSC, CLARIN-ERIC, DARIAH-ERIC, ECCSEL-ERIC, EMBRC-ERIC, Euro Bioimaging-ERIC, Lifewatch-ERIC, BBMRI-ERIC, ECRIN-ERIC, CLARIN-ERIC, PHENOME-ESFRI, EMPHASIS-ESFRI, ELIXIR-ERIC, COESO, DISSCO
- Various stakeholders representing operational level: Univ. Stockholm, Univ. of Ljubjana, Univ. of the Basque Country, Sant'Anna School of Advanced Studies, French National Pilot Committee for Digital Ethics, Helmholtz Association, EARMA-ERION, OpenAIRE, PANELFIT project.

Significant observations from the first workshop have been incorporated in the preparation of the second workshop:

- The need to differentiate generic infrastructures that are dedicated to OS, on the one hand, and on the other hand thematically focused infrastructures committed to OS, more "in process" of developing OS policies or contributing to their development;
- Field-based and national legislation-based differences between infrastructures;
- The fact that technical issues are not the only challenges to implement responsible OS (or even not recognized as major by the infrastructures themselves).

Some findings and issues from the preliminary report were identified for further discussion in the second series of workshops:

- Research infrastructures are in a position to raise awareness, offer training, elaborate guidelines and practical tools.

[8] See the extensive list in appendix 1

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

- Guidelines and training are essential, but they must be discipline-specific and/or area-specific in order to be considered relevant.
- Some ERICs maintain legal information platforms for their communities: they face difficulties to keep these platforms updated, especially with the diversity of member states' legal regulations of intellectual property. Different national traditions and practices between EU member states sometimes prevent ERICs and European infrastructures from establishing common guidelines at the European level.
- There is a need to help researchers to distinguish between the legal basis (personal data) & the fundamental ethical requirements.
- There are needs for education and tools in order to avoid the perception of these tools as undue bureaucracy, and therefore be overlooked or addressed a-minima and/or irrelevantly and inconsistently.
- There is a need to clarify the responsibility of institutions that employ researchers, and/or responsibilities which can only be enforced by national policies in each member state.

The expectation of obtaining the views of different fields of research was examined, through the differentiation between scientific disciplines offered by the Frascati Manual of the Organisation for Economic Co-operation and Development (OECD, 2015[9]). It differentiates between six fields of research and development (so-called broad classification), each of which contains several subfields (so-called second-level classification). However, this classification does not fit very well with the broader spectrum of research infrastructures and presents the risk of going into too many details for the workshop format. An appropriate option was not to focus on the discipline but on the overall type of practice / methodology of research. These three families of scientific methodological practices (see below) are not exclusive from one another and all disciplines more or less rely on these practices. A given discipline, whatever their broader field, can be represented as performing a combination of these three methodological practices, defining a volume in a three-dimensional space where these practices are the dimensions. These families of practices differ in their relationship with the research variables and outputs outlined above and therefore in the spectrum of their strengths and weaknesses in terms for instance of research integrity.

We used the following typology of methodological practices:

- **Observational**. Understanding of a natural, social or human phenomenon, without any control of the operator or any ability to repeat the observation. Describing the conditions of the observation is crucial, including the date since time is an unrepeatable variable. This is for instance largely the case of climatology or biodiversity studies, or of many sociological studies.
- **Experimental**. Understanding of a phenomenon in a system of which some or all parameters are controlled by the researcher. Repeatability is possible, and crucial. This is for instance largely the case in biomedical disciplines. Simulation approaches can be assimilated here.
- **Theoretical**. Understanding of a phenomenon with no or limited connection with reality at the time of the analysis, although reality may be inspirational and return to observation after the

[9] OECD (2015): Frascati Manual 2015: *Guidelines for Collecting and Reporting Data on Research and Experimental Development, the Measurement of Scientific, Technological and Innovation Activities*. OECD Publishing, Paris. https://doi.org/10.1787/9789264239012-en

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

theoretical phase can be envisaged. Describing the hypotheses, proofs, narratives, interpretations, models, arguments etc., the reasoning and its outcomes are crucial.

The opportunity of broadening the discussion on other productions than data was taken through questions on all research productions, asking for concrete examples, resources and use cases. Although not all responses can be used directly in the mapping and analysis, they can be considered as supplementary materials towards the Knowledge Hub, to answer the need for an in-depth analysis of the challenges and some use cases, tool collections.

For the second session of workshops, three sessions were organized by type of research practice in 2022, on the 6th of October, 25th of October and 9th of November, in addition to one open discussion with representatives of citizen science projects or platforms on the 13th of July during the cross-SWAFS forum.

The goals of these workshops, based on the preliminary report, were to get:

- An end-user perspective on the provision of "responsibility" certification services by OS infrastructures;
- A consideration of the cost of development and of maintenance of such specialized services for OS infrastructures;
- A documentation of the question of informed consent, in various disciplinary or thematic contexts;
- A mapping of the challenges, current solutions, gaps and expectations across disciplines, thematic fields and geographical situations.
- The possibility that data management plans can be efficient vectors of responsible science principles in practice, to be valued by the ROSiE Knowledge Hub.

In order to prepare these goals, a list of questions was sent to participants in advance, but it was agreed that they could digress from the list of questions to address the points that seemed essential to them, from a practice-oriented point of view[10].

For all the four workshops, we issued a verbatim or a transcript from which we conducted a thematic qualitative analysis described below, by major themes addressed by the participants.

[10] See appendix 2

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

3 Workshops results

3.1 Challenges for Open Science Infrastructures

Participants to the workshops shared several challenges encountered by OS Infrastructures (OSI). We classified those challenges in three main categories, according to the ROSiE research objectives: 1) challenges mainly related to research ethics (RE); 2) challenges mainly related to research integrity (RI) and 3) challenges mainly related to the FAIR principles. All challenges have RI implications and several challenges are interrelated. (See figure below)

3.1.1 Challenges related to research ethics

The challenges mainly related to research ethics (i.e., related to data collected on human subjects) were:

- ✓ **Incidental Findings** can be a challenge for OSI – namely, to find unexpected results and observations in a set of data.

"*The whole principle of open data is that you get unexpected results from open data, as people see new patterns that you don't necessarily see in the first instance. So, giving a consideration on how you do that for really successful citizen science projects is absolutely crucial in my mind*[11]*."*

- ✓ **Informed Consent**: an adequate informed consent (i.e., considered as such by traditional evaluation of research ethics committee) can be challenged by certain OS methodologies. To be transparent enough about all particular uses and re-uses in the context of OS is not always possible, simply because they (especially the re-uses) cannot be envisioned at the time of data collection: in this case, the consent cannot be considered entirely "informed" over the course of the data life cycle.

"*Our ethics committees are not equipped to help us with such complicated cases and methodology*"

"*The complaint about broad consent had been that it kind of limits the tissue donors' autonomy because the particular uses of the data is not transparent to them*"

- ✓ **Privacy & Confidentiality** requirements are a challenge for OSI, namely in limiting the access to data (especially sensitive and personal data). OSI often confront a dilemma between the level of anonymization of open access data and the level of utility and reusability. Privacy and confidentiality requirements may also impair the utility of some support - for example, when limiting the sharing of some information in DMPs:

[11] All quotes provide from the discussions in the workshops.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

"*I searched for DMP models in order to benchmark how others had addressed these issues. Another important feature [of data sets] is that they often contain private information that cannot be made public. The available DMPs are not always super useful*[12]."

This challenge is highly dependent on the nature of data (i.e., sensitive or personal data, digital or physical data) and thus indirectly dependant on the discipline involved. Those limitations are legally grounded in Europe – e.g., with GDPR requirements.

"*I think the tagline is 'as open as possible and as legally as necessary'*"

Privacy and Confidentiality may thus limit data accessibility and reusability if not properly addressed ab initio, which they cannot always be.

3.1.2 Challenges related to research integrity

- ✓ **Acknowledgment of Contribution & Intellectual Property:** OS may create situations where the contribution of the person who collected the data may not be recognized in someone else's discovery, made using this data:

"*There's always this hesitation: that releasing the data, somebody might discover something really interesting and then, you know, you put all that effort in and somebody else gets the Nobel Prize."*

Although this challenge is not specific to OS, the OS context (namely, wider dissemination of data, diversity of possible actors having access to it etc.) can increase the impression of a conducive environment for this issue to occur. This perception is important to be aware of, as it may constitute an obstacle to the opening of data.

Sometimes, what is difficult is to give the appropriate credit to all individuals that contributed to the data collection, for example for historical data:

"*So usually, let's say during a colonial expedition, the scientist were primarily the European scientists, when a lot of local helpers and local experts and local scientific helped them to collect the data. And they're usually not mentioned. […] So how do you connect those things and find those persons and provide attribution to and give them the credit to that?"*

Data curation is one of the 14 types of contribution to research recognized in the CRediT taxonomy[13], but a formalized system is missing to acknowledge the contribution of 'data curators' for instance when citing a paper: they are at the best mentioned in a footnote or in the 'Acknowledgment' section.

- ✓ **Conflict of Interests:** while not explicitly discussed, this challenge was mentioned in the context of research equipment cofounded by private and public institutions.

- ✓ **Data Quality & Integrity** refers to the question how to guarantee the quality of the data, especially for citizen science data platform (data not collected by professional scientists):

[12] Translated in English by the authors of the report.
[13] https://credit.niso.org/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

"*You have identified a fundamental issue about how you guarantee that the quality of the data is high. And it is an issue that we, as a citizen science community, are going to have to solve if we want to see the scale of our activities growing."*

Part of this challenge is related to data collection and dissemination: to limit missing information and human errors (more likely to occur with a high number of users) and ensure that the data was not fabricated or falsified. Ensuring data quality and integrity is not only a challenge for OS, but OS practices amplify the impact of a lack of quality or integrity by collecting and disseminating data at large – even inaccurate or otherwise unsuitable data.

3.1.3 Challenges related to some of the FAIR principles

- ✓ **Accessibility**: data ownership may first challenge accessibility. Data does not necessarily belong to the hosting platforms. It can also belong (depending on the nature of data) to the person, to the institutions or "projects", which can complicate the possibility of free circulation and control of OSI on these data. In the specific case of sharing research papers[14], disciplinary traditions of authorship or copyright may limit open access, for example when visual elements are part of the data to be shared:

"*Using images or video to support a discussion implies questions of copyright. Very often, in disciplines like history of art, free circulation of articles and publications is restricted, or it would imply removing images that are fully part of the scientific argument.*[15]"

The difficulty for OSI to secure long-term funding can also hamper open access: the underlying question being how to reconcile the requirements of open access with the need for incomes (for example, editor's incomes). The question of the possibility for open data to be freely accessible (without charging users) has been raised by some participants of the workshops. Access limitation may also differ according to the origin of the funds (public or private) and the specific requirements associated with their use (for example, geographical restriction). A need for guidance and policy in order to ensure the appropriate use of public fund in this situation was raised:

"*We have some right concept […] and some proper policies that ensure that we don't waste taxpayers' money but that we have it used according to the aims that we have and towards what we are supposed to deliver for getting this money that we have from the countries from commission and ultimately from taxpayers. So for me, this is still also a point to consider when you talk about responsible Open Science*".

Finally, some ethical and legal requirements may challenge open access (for example, for sensitive data, see Privacy and Confidentiality).

- ✓ **Interoperability** can raise challenges in a different way, because of the diversity of the disciplines, of the platforms and of the way they are managing data, of practices or even

[14] Research papers may be considered as data, especially for Humanities and social sciences.
[15] Translated in English by the authors of the report.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

because of cultural diversity. Participants acknowledged the difficulty to create a template that "works for everything", and the challenge to adapt OSI to the requests of researchers from different domains.

"*I mean, we are dealing with a similar situation and our extra challenges is that we were working with 270 different museums across 23 different countries. So everybody has their own local practices and then we have to come up with the European practice and sometimes they don't match*."

"*So the main problem there is, as I understand it, is if there was unity of the regulations across countries, there are also some cultural differences. Then how do you implement everything into an infrastructure research that tries to enable some transnational work?*"

This challenge is strongly interrelated to the re-use challenge.

- ✓ **Re-use:** while it was recognised that one of the aims of OS is to allow the re-use of data and methods to make results replicable, several participants highlighted challenges in ensuring the re-use of data, including (1) the lack of harmonisation in the way data is collected and stored, (2) the difficulty to ensure data relevance over time and (3) the difficulty to sharing qualitative data in a re-usable format. Another challenge is to ensure that re-use would benefit the communities that initially created the data, or at least that they receive a fair credit for their collective effort.

"*So, making it available, you know, more generally, even not to the public, but just to a wider scientific community is not only difficult but also maybe not so meaningful because nobody else besides the people who actually gather the data, know how to analyse it."*

"*Originally, sometimes, they were written for microbes (for instance). And so people working on plants or animals were interpreting as they wanted and it makes them very heterogeneous in the end on the dataset*."

- ✓ **The Digital Divide** refers to the risk that OS, in a counterintuitive manner, would favour wealthier countries and institutions, or ones that already have some advantage over others. This challenge affects all the previous challenges related to the FAIR principles. Imbalance of resources between countries and between institutions or even people in the same country may challenge a genuine open access, as well as findability, interoperability and reuse of open data. High-quality data may not be accessible for all, or intelligible or usable to all in an equal manner. Less privileged institutions may not have the tools, resources and/or skills to use open data. A need to make OS more inclusive has been underlined.

"*If you have metadata that, for instance, contains references on specific lab equipment that is used to create that data, and then one is state-of-the-art and another is not seen as state-of-the-art, then this devalues the data automatically through the metadata, and it's again a question of resources that some have and some don't.*"

"*I was just thinking about […], you know, partnerships that you have within Europe... Do they tend to be, you know […] researchers from resource-rich institutions or do you feel that that's really that your resources so therefore anyone to use […], also from resource-poor areas."*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

3.2 Some thoughts on policy and guidance

During the workshops, several discussions started about benefits and gaps of current policies and guidance (mainly, discussion about Data Management Plan was encouraged). A lack of policies specifically dedicated to responsible OS was underlined.

- ✓ **Data Management Plan (DMP)** is sometimes considered unclear or difficult to implement, for both researchers and citizens (in the context of citizen science). Participants acknowledged a need for support and training for researchers on how to use DMPs.

"*One of the issues about the DMP is that people don't have the training to actually properly develop them or implement them.*"

For others, DMPs could also be seen as a useful tool, a way to reflect on open data ethical considerations[16] and even a good way to standardize data collection and storage. Thus, DMPs risk being perceived as a mere burdensome bureaucratic exercise if not used in the right way.

- ✓ **Diffuseness**: OSI and platforms have the potential to bring together data collected in many different countries and to provide access to this data to individuals around the world. These countries do not have the same jurisdiction or data protection laws, which complicates the possibility of Accessibility and Reuse.

3.3 Recommendations

3.3.1 General Recommendations

Participant formulated some general recommendations for the development of responsible OSI:

- ✓ Training should be developed (both for users and for platform makers);

"*Yeah, on helping researchers integrating this whole idea of FAIR data and OS into their real activities when they perform research, beyond a single experiment*."

- ✓ Particular attention has to be paid to the plurality of disciplines and issues related to OS, i.e., OS should be thought of in a "*cross-disciplinary manner*";
- ✓ More resources should be dedicated to OS;

[16] Some discussion about GDPR also go in that sense: as mandatory, it could be seen as a way to engage reflection on data management.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

- ✓ The CARE principles should be added to the FAIR principles in order to better take into account the human that generated the data[17] ;
- ✓ When guidelines on responsible OS are developed, they should target specific aspects (e.g. data collection, consent forms, legal aspects, etc.), and they should be operational, up-to-date and ready to use materials.
- ✓ Responsibilities should be distributed and infrastructures should collaborate.

Participants recognised that responsible OS is only possible on the basis of a shared responsibility. They stressed that platforms and infrastructures do not have the same roles and therefore not the same responsibilities. For example, regarding data quality and integrity, while infrastructures are responsible for being transparent about their processes in order to allow users to assess themselves the data quality, platforms would have a role to play in managing data or guide data management. Platforms participate in the responsibility of data integrity, but they are not the only one responsible: owner(s) of the data, but also scientists and citizens who collect it, have their role to play. A participant mentioned that it is not possible for a single infrastructure to meet all the challenges of responsible OS: this requires the collaborative work of several infrastructures. The DMP was recognized as a good way of determining who is responsible for what, and to do so upstream of the data collection.

It should be noticed that the participants did not mention technical issues or solutions either spontaneously. When they were specifically questioned on these matters, technical options did not seem to be a priority in addressing the above-mentioned challenges.

3.3.2 Recommendations for the ROSiE Knowledge Hub

- ✓ Propose a “train the trainer” format for educational tools;
- ✓ Develop use cases as helpful support for reflexion of end-users;
- ✓ Disseminate and implement resources in a way that create “commonality” across the global community;
- ✓ Create a network between communities across countries;
- ✓ Develop embedded support into the DMP with concrete examples, or even a set of (ethical) questions to encourage reflexivity.

[17] The CARE principles (Collective benefits, Authority to control, Responsibility, Ethics) were developed by the Global Indigenous Data Alliance in order to ensure better inclusivity and better consideration to indigenous communities in open science and data management.
See: https://www.gida-global.org/care

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

3.4 Potential Solutions

Participants formulated several potential solutions to address the challenges previously identified and allow responsible OSI. Potential solutions are concrete mechanisms to develop, and often rely on real-life experience and examples from participants' practices.

- ✓ **Building Integrated Systems** that allow harmonization of data storage, organisation and evaluation of quality. This may especially help to address re-use and Interoperability challenges.
- ✓ **Creating Common Charters:** the creation of a charter to be adopted by all the actors collecting and sharing the data into a same infrastructure or platform[18]. This common charter may help address the challenges mentioned above.
- ✓ **Implementing Data Justice Plan** as worded by a participant:

"*One of the things that we've come up with in that work is that, rather than just having data management plans which look at the data life cycle throughout the project, that are mostly focused internally, we should supplement those with something that we call a data justice plan which would allow for projects to really showcase, and then track how the data benefits the citizens and what sort of citizens contribute. And what is the benefit of our data for the community, for the citizens / for the participants, and have that documented, alongside data management, could be part of a data management plan.*"

This echoes the CARE principles previously mentioned and may help address the digital divide challenge and the reuse challenge, among others.

- ✓ **Creating Playbooks:** short and easy playbooks to help understand how to create a DMP[19].
- ✓ **Rewording Consent Forms** to adapt them to the context of OS, thus addressing the challenge of Informed Consent[20].
- ✓ **Implementing an OSI Service Level Agreement** to be adopted by platforms, that describe expectations, procedures and guidelines of the OSI and the data gathered – for example, related to the FAIR principles or on how platforms will work together. Such agreement may help address all the challenges presented above.
- ✓ **Developing Wikidata Models** refers to the development of an expertise community annotation and curation process to address data quality & integrity and ensure trust.

"*And in one, one way we're also looking into kind of the wikidata model […], where you could go and edit data or upload data. So in particular, we are thinking about this as a community curation and community annotation model. [...] So we want to bring that sort of, you know, community level where different expertise can come in and annotate or provide input on the quality of the data*."

[18] See for example : https://rdmkit.elixir-europe.org/
[19] See for example : https://cetaf.org/resources/best-practices/
[20] See for example : https://elixir-europe.org/about-us/how-funded/eu-projects/converge

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

This goes in combination with a Flag and Annotation System (i.e., to identify errors, missing info etc.) which can lead to a "data quality score or trust score". This model can work with the implementation of a Persistent Expertise Profile (a guarantee of the credibility of the user and therefore of the quality of the flag and annotations it generates) that allows the origin of each piece of data, annotation or flag to be informed.

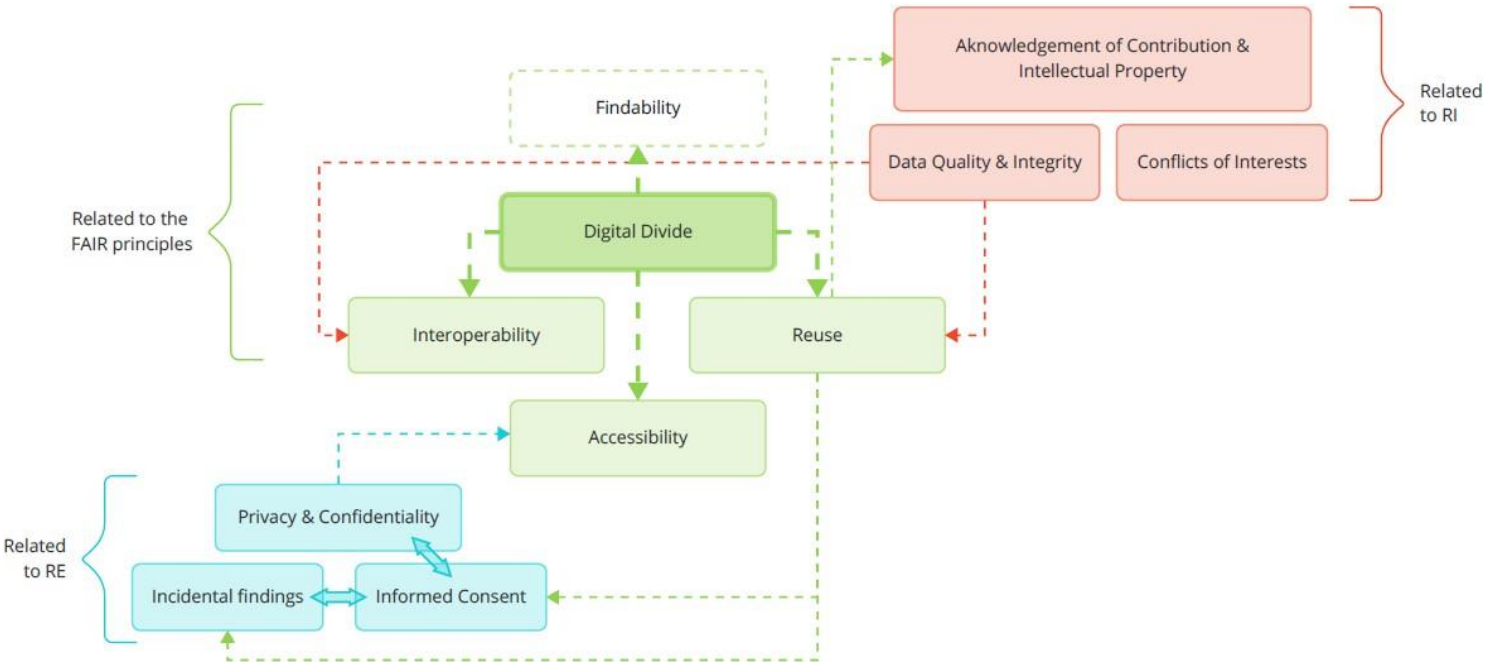
"*And of course we also try to show the value. So for example, we spend a lot of time on thinking about persistent identifiers and metadata and repositories so we can show the researchers, "okay, if your data is in a trusted repository with a persistent identifier that's linked to your publication, you know, others can find it, others can cite it". So then you can you can sort of think about the value of it. So sort of the carrot*."

- ✓ The persistent expertise profile can also help meet the challenge of contribution recognition.
- ✓ **Lowing Data "Quality":** some offered to use lower quality data for better inclusivity (addressing thus the Digital Divide challenge). For example, it may involve the use of XML-TEI[21] encoding or the use of platforms that allow the adjustment of video bandwidth according to the quality of the users' connection.

In conclusion, the main challenges mentioned by the participants to the workshops can be summarized by this visual representation:

[21] For more information, see: https://tei-c.org/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430



3.5 Ways forward

During the workshops, participants mentioned several case examples of good practice and tools. Although these do not necessary belong to this report, they can constitute an interesting supplementary material towards the preparation of the knowledge hub.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

4 REFERENCES

4.1 Reports & publications

Armeni, K et al., “Towards wide-scale adoption of open science practices. The role of open science communities, 2021.

Carvalho, A.S. & Strecht Almeida, M. (2022): D.1.1: *Report on the relationship (tensions, challenges, overlaps) between RI, the wider RE perspective and OS*. ROSiE: Responsible Open Science in Europe, unpublished project deliverable.

D’Andrea, L., Berliri, M. and Federico M. (2018). Fostering improved training tools for Responsible Research and Innovation, Summary Report, Deliverable 1.3.
Available: https://zenodo.org/record/1434355.

Le Gall, O. & Medves, M. (2021): D 6.1: *Preliminary analysis and mapping of existing European and national Open Science infrastructures with regard to promoting responsible Open Science*: Responsible Open Science in Europe, unpublished project deliverable.

Lindemann, T. (2021): D.3.1: *Report on a strategy to engage Stakeholders*. ROSiE: Responsible Open Science in Europe.
Available : https://rosie-project.eu/deliverables/

Lindemann, T.; Häberlein, L. & Hövel, P. (2022): Report on interviews. ROSiE: Responsible Open Science in Europe.
Available : https://rosie-project.eu/deliverables/

Maire, A. The digital disruption of science: Governments and scientists toward an “Open Science”. The Digital Transformation of Labor, 2019.

Martin, E. G., & Begany, G. M. (2017). Opening government health data to the public: benefits, challenges, and lessons learned from early innovators. Journal of the American Medical Informatics Association, 24(2), 345-351.

Mežinska, S.; Mileiko, I.; Neiders, I. & Kalēja, J. (2022): D.2: *Report on social challenges and implications related to Open Science*. ROSiE: Responsible Open Science in Europe, unpublished project deliverable.

OECD (2015): Frascati Manual 2015: *Guidelines for Collecting and Reporting Data on Research and Experimental Development, the Measurement of Scientific, Technological and Innovation Activities*. OECD Publishing, Paris. https://doi.org/10.1787/9789264239012-en

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

4.2 Website contents and resources

Consortium of European Taxonomic Facilities (CETAF), resources portal : https://cetaf.org/resources/best-practices/
(last consultation : 31.01.2023)

Elixir Europe, CONVERGE project : https://elixir-europe.org/about-us/how-funded/eu-projects/converge
(last consultation : 31.01.2023)

Elixir Europe, The Research Data Management toolkit for Life Sciences : https://rdmkit.elixir-europe.org/
(last consultation : 31.01.2023)

Global Indigenous Data Alliance (GIDA), The CARE principles (Collective benefits, Authority to control, Responsibility, Ethics) : https://www.gida-global.org/care
(last consultation : 31.01.2023)

European Open Science Cloud (EOSC) Portal : https://eosc-portal.eu/
(last consultation : 31.01.2023)

National Information Standarts Organization (NISO), Contributor Role Taxonomy CRediT : https://credit.niso.org/
(last consultation : 31.01.2023)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

5 Appendix

5.1 Appendix 1 – List of participants (2021 & 2022 workshops)

- EOSC:
  - Karel Luyben (president of EOSC)

- ROSiE:
  - Heidi Beate Bentzen (WP2)
  - Carole Chapin (WP6; organizer)
  - Rosemarie de la Cruz Bernabe (WP1, WP9)
  - Lisa Häberlein (WP3)
  - Arild Jansen (WP6)
  - Panagiotis Kavouras (WP6, WP8)
  - Keziah Chanyisa Khayadi Dash (WP9)
  - Teodora Konach (WP5)
  - Olivier Le Gall (WP6; organizer)
  - Tom Lindemann (WP3)
  - Maud Medves (WP6; organizer)
  - Signe Mežinska (WP7)
  - Vivian Nchanchou Mbanya (WP9)
  - Kadri Simm (WP1)
  - Lisa Tambornino (WP3)
  - Mariana Vidal Merino (WP4)
  - Nathalie Voarino (WP6)

- ERICs / ESFRI :
  - CLARIN-ERIC: Paweł Kamocki
  - DARIAH-ERIC: Erszébet Toth Czifra
  - DISSCO-RI : Sharif Islam
  - ECCSEL-ERIC: Sébastien Dupraz
  - ELIXIR-FR : Anne-Françoise Adam-Blondon
  - EMBRC-ERIC: Gemma Gimenez Papiol
  - EMPHASIS-ESFRI : Sven Fahrner, Jacques le Gouis (PHENOME-FR)
  - Euro Bioimaging-ERIC: Aastha Mathur
  - Lifewatch-ERIC: Christos Arvanitidis
  - BBMRI-ERIC: Michaela Theresia Mayrhofer, Mónica Cano Abadia, Ilaria Anna Colussi

- Other participants
  - Jonas Åkerman (Univ. Stockholm & EARMA-ERION)
  - Denise Amram (Sant'Anna School of Advanced Studies & PANELFIT project)
  - Ilmari Jauhiainen (TSV – Federation of Finnish Learned Societies)
  - Emmi Kaaya (Univ. Tartu)
  - Claude Kirchner (French National Pilot Committee for Digital Ethics)
  - Mojca Kotar (Univ. Ljubjana & OpenAIRE)
  - Hanna Lahdenperä (TSV- Federation of Finnish Learned Societies)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

- Iñigo de Miguel Beriain (University of the Basque Country & Project PANELFIT)
- Antonia Schrader (Helmholtz Association)
- Alessia Smaniotto (OPERAS ; Coeso project)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

5.2 Appendix 2 – Questions for the 2022 workshops

1. Have you identified challenges related to responsible open science in your activities or in the activities of your stakeholders?
2. Have you experienced existing technologies that safeguard responsible OS in the context of experimental practices in research?
3. What do you think would be end-user perspective on the provision of "responsibility" certification services by open science infrastructures?
    a. In other terms, how to ensure end-(re)users they can trust the data set (e.g. that it has not been falsified or fabricated), or open access research tools, or any other type of FAIR open resources that they are about to re-use?
4. Do you think that data management plans can be efficient vectors of responsible science principles in practice, to be valued by the ROSiE knowledge hub?
5. Do you consider the repository/infrastructure you represent a data controller, a joint data controller, or a data processor?
6. What are your data access procedures?
7. What are your procedures for ensuring that the reuse of data is within the limits of the original consent to research participation provided by the research participants?
8. What are your procedures for ensuring that the data reuser has a lawful basis for data processing according to the GDPR?
9. What are your procedures for ascertaining the identifiability of the data set?
    a. in other words, whether the data is anonymous or if it contains any personal data according to the definition in the GDPR ("any information relating to an identified or identifiable natural person")?
10. Does the research institution depositing data in your repository have any say regarding any of the repository's technical and organizational measures? If yes, which technical and organizational measures?

Additional questions

11. Do you face difficulties or have you implemented solutions about ethical or social challenges of opening science?
12. Which procedures are in place for informing initial research participants of findings resulting from the reuse of data about them that are of relevance, for instance, to their health or other personal concerns?
13. Why do you think more private companies do not currently make data available for reuse? How can we increase the number of private companies that make data sets available for reuse through repositories?
14. Scientific research is a global endeavor, but transfer of personal data to researchers outside the European Economic Area, and provision of remote access to researchers outside the European Economic Area, is subject to strict restrictions in the GDPR. What are your procedures for data transfers to outside the European Economic Area? (and symmetrically, for data transfers from outside the EEA?)
15. If a research participant wants personal data about them to be modified or erased, what are your procedures for achieving this?

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430