

SCIENCE EUROPE

PRACTICAL GUIDE TO SUSTAINABLE RESEARCH DATA

Maturity Matrices for Research Funding Organisations,
Research Performing Organisations, and
Research Data Infrastructures





June 2021

Practical Guide to Sustainable Research Data - Maturity Matrices for Research Funding Organisations, Research Performing Organisations, and Research Data Infrastructures

DOI: [10.5281/zenodo.4769703](https://doi.org/10.5281/zenodo.4769703)

Authors: Tommaso Boccali (National Institute for Nuclear Physics), Anne Elisabeth Sølsnes (Research Council of Norway), Mark Thorley (UK Research & Innovation), Stefan Winkler-Nees (German Research Foundation) and Marie Timmermann (Science Europe).

Acknowledgments: The authors would also like to thank the members of the Science Europe Working Group on Data Sharing and Supporting Infrastructures, in particular further members of the task group 'Sustainable Research Data' and the participants of the Science Europe workshop 'Achieving Sustainable Research Data' (January 2021).

Editing and Review: Giorgia Battiatto, Iwan Groeneveld, Lorna Stokes, and Lidia Borrell-Damián (Science Europe).

For further information please contact the Science Europe Office: office@scienceeurope.org

© Copyright Science Europe 2021. This work is licensed under a Creative Commons Attribution 4.0 International Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited, with the exception of logos and any other content marked with a separate copyright notice. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/> or send a letter to Creative Commons, 444 Castro Street, Suite 900, Mountain View, California, 94041, USA.

Table of contents

Foreword by Professor Roland Fischer and Professor Melanie Welham	4
Introduction	6
How to use the maturity matrices?	11
Research Funding Organisations (RFOs): Maturity matrix for sustainable research data	19
Research Performing Organisations (RPOs): Maturity matrix for sustainable research data	29
Research Data Infrastructures (RDI): Maturity matrices for sustainable research data	39
References	46



Foreword by Members of the Science Europe Governing Board

Professor Roland Fischer,

Vice-President of German Research Foundation (DFG), and

Professor Melanie Welham,

Executive Chair, UK Research and Innovation (UKRI)

Open Science builds on the fundamental features of research and innovation: transparency, openness, verification and reproducibility. It is one of the main elements in Science Europe's strategy for 2021–2026, and together with its peers from the Science Europe membership, UK Research and Innovation (UKRI) and the German Research Foundation (DFG) support the aim to deeply embed it within the practice of research. To achieve such change, appropriate policies and practices must be in place throughout the research and innovation system.

One of the key elements of Open Science is the availability of research data. Research data are an essential foundation for scientific work; their diversity reflects the wide range of scientific disciplines, and of research interests and methods. Sharing research data for re-use to support scientific progress is increasingly becoming the norm and a growing number of organisations already expect and foster an open behaviour regarding such data.

This Science Europe Practical Guide aims to provide guidance to both Science Europe Member Organisations and other interested research stakeholders on how to ensure the long-term preservation and accessibility of research data, offering a framework for researchers to share their outputs in a sustainable way. This directly contributes to the traceability and quality of scientific work and enables researchers to carry on work begun by others.

The matrices presented in this guide can support organisations to develop and enhance their own policies and practices towards the sustainability and openness of research data, and encourage dialogue and collaboration with like-minded organisations. This directly supports the advancement of Open Science as a whole and helps it become an increasing part of a shared research culture.

Science Europe's Member Organisations and their peers in the research and innovation landscape play an essential role in the development and implementation of Open Science policies, both individually and collectively. We hope that this guide will be a useful tool to advance in this ambitious and important endeavour.

Professor
Roland Fischer

Professor
Melanie Welham



Introduction

ON THE WAY TO MAKING OPEN SCIENCE THE ‘NEW NORMAL’

The research and innovation (R&I) system is undergoing a paradigm shift that aims to make scientific knowledge accessible for all by promoting the unhindered dissemination of research outputs. Today, many collaborative initiatives promote Open Science as the ‘new normal’ and progressively define what a modern research system should look like. Such initiatives include the European Open Science Cloud (EOSC), cOAlition S, the San Francisco Declaration on Research Assessment (DORA), and OA2020. Open Science is also embraced by Horizon Europe, the European Commission’s new Framework Programme for Research and Innovation.

Sustainability of research data¹ refers to the long-term preservation and accessibility of data collected and produced in research processes. It is essential for the reproducibility and re-use of these data that they be shared according to the principle ‘as open as possible, as closed as necessary’, the FAIR principles², and the TRUST principles.³ Ensuring the sustainability and interoperability of research data is an indispensable element to make Open Science the normal way of doing research, and is key for the successful implementation of the EOSC.

Researchers generate considerable amounts of data through their projects, which can be funded by various sources and supported by different home institutions. Research Data Management (RDM) is an essential element of the research process. It is an ongoing activity throughout the data lifecycle, including the active organisation

and maintenance of data during the research process and suitable archiving for future re-use. It requires researchers and their host institutions to consider, from the start of a research process, how to collect, re-use, store, and curate their data.

Research Funding Organisations (RFOs), Research Performing Organisations (RPOs), and Research Data Infrastructures (RDIs)⁴ play an important role in the advancement of Open Science and RDM. These organisations need comprehensive strategies that include a strong commitment to RDM, referred to in this document as an ‘agenda for research data’, to set the frameworks that enable efficient data sharing for the long term. Their agenda for research data should include several strategic areas, explained in the following sections of this guidance, that need to be addressed to make research data sustainable.

Effective collaboration with other actors in the R&I system and aligned agendas for research data contribute to fulfilling this aim. Achieving sustainable research data is a journey that must be jointly undertaken by all actors involved while allowing each one to travel at their own pace and choose their final destination according to their own strategic priorities and needs.

THE TOOL: COMPLEMENTARY MATURITY MATRICES

Science Europe aims to provide strategic guidance to its Member Organisations and the broader research community on the necessary steps to help researchers carry out their work in the best conditions. Areas of action include financial investment, policy development and implementation, infrastructure availability, career development and skills (both for researchers and provision of data scientists), and clearly assigned roles and responsibilities among the different actors.



This guide to sustainable research data contains three complementary maturity matrices: one for RFOs, one for RPOs, and one for RDIs. Maturity matrices are a tool to assess how well developed a particular area is and to identify what next progression steps look like. They were developed by experts from Science Europe Member Organisations represented in the Working Group on Data Sharing and Supporting Infrastructures (WG DSSI). Further input was sought from RFOs, RPOs, universities, and RDIs across Europe during a validation workshop in January 2021.⁵

The three maturity matrices presented in this guide complement each other to support the alignment of policies and requirements for data sharing across organisations, as Science Europe has promoted for many years.⁶ When taken up by the research community, they will make it easier for researchers to comply with RDM requirements when working in different organisations, with diverse research funding organisations, or in varying national contexts.

SUSTAINABILITY OF RESEARCH DATA: A JOINT RESPONSIBILITY

Sharing and re-using data requires organisational commitment from RFOs, RPOs, and RDIs, as well as resources and skills at individual, organisational, and government level to make sure the necessary funding and structures are available. Policies need to be developed in a timely manner for existing data infrastructures and considered when new infrastructures are developed and set up.

RFOs, RPOs, and RDIs should aim for more concerted data sharing approaches to rationalise their efforts and increase the usefulness and efficiency of RDM for researchers. A coherent sustainable framework for data sharing and interoperability is most needed. When developing, updating, and implementing their policies and practices, organisations should take relevant international developments into account.

Examples are the establishment of the EOSC, which will play an important standard-setting role, or legislative initiatives by the European Commission based on its 2020 European Strategy for Data.⁷

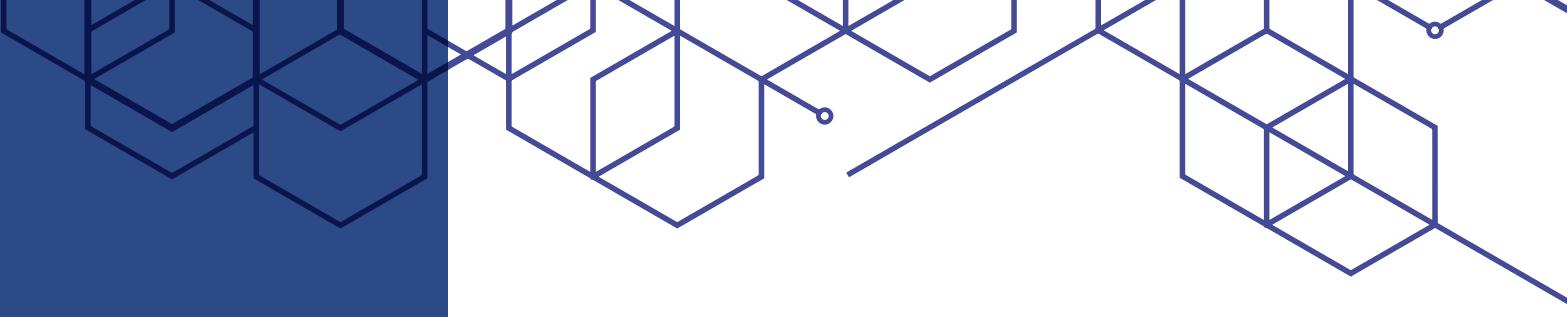
This publication provides guidance tailored to the three organisation types on how to develop and/or adapt their policies and practices. As the matrices are complementary, organisations are encouraged to also use the matrices for discussion and collaboration with other organisations to align approaches and define each organisation's role in their national or discipline-specific context.

By developing its own organisational agenda for research data, every organisation will define the role it will play to make research data sustainable and will be able to better contribute to the alignment of approaches among organisations.



SUSTAINABLE RESEARCH DATA

How to use the maturity matrices?



Goal-setting Depending on Strategic Priorities, Missions, Mandates, and Needs

This guidance is designed to support RPOs, RFOs, and RDIs in developing their agenda for research data to achieve sustainable data sharing and interoperable systems. It takes the form of three complementary maturity matrices to allow collaboration with other organisations.

The matrices present a framework and propose actions in six essential areas:

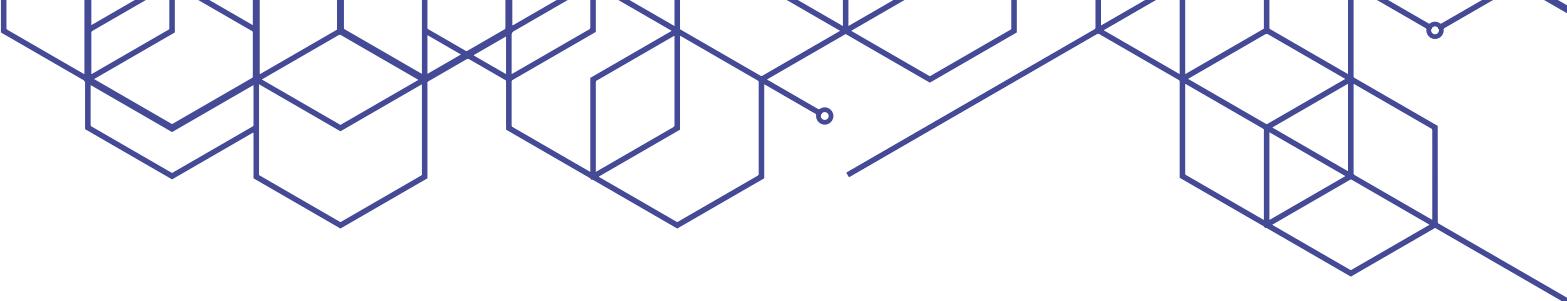
- Organisational engagement and commitment
- Policy environment
- Financial aspects
- Training
- Technical preparedness
- Communication and awareness raising

These areas were defined following discussions among the experts from the Science Europe's WG DSSI, based on their experience and expertise, and extensive desk research. The areas were discussed and validated both by the WG DSSI and during a validation workshop with external experts.

The matrices allow organisations to assess their own situation in relation to the six areas and to plan their next steps according to their organisational strategic priorities, mission, mandates, and needs. The matrices can also be used to compare actions, set collaborations, and/or seek and monitor alignment with other organisations. The interpretation and the application of the matrices may vary depending on internal organisational policies and needs, and/or on external factors ruling data management.

To assess its state of development towards sustainable research data, an organisation needs to consider its respective matrix as a whole. It should be seen as a guide for progressive development, starting with the first step for each area, ideally completing all actions proposed under one step before moving to the next one. It is, however, acknowledged that there will be cases of organisations that have reached higher progression steps without having completed all actions of the previous steps.

Organisations might find that they are more advanced in certain areas than in others. Depending on their organisational mission, strategic goals, and mandate organisations will need to define which level they want to achieve in any given area. Not every organisation will have to reach the highest maturity level in all areas as certain actions may not be within the remit of their mission or mandate. Therefore, the potential next step needs to be defined for each area individually and in line with each organisation's strategic goals.⁸ Some organisations may take on a driving role in a specific area, setting standards and leading policy developments, while others might have different priorities and could contribute to these efforts with expertise without having to take the lead. As the level of maturity advances, the level of collaboration with other RFOs, RPOs, RDIs, or other stakeholders will also increase.



Definitions of progression steps in the matrices

The matrices present three progression steps for each of the six key areas:

- **Plans to develop:** The organisation has acknowledged the need to take action in a given area and is developing/has developed plans on how to proceed.
- **Development ongoing:** The organisation has done the groundwork in a given area to achieve the sustainability of research data, though more refinement is needed.
- **Developed on organisational level:** The respective area is addressed on a mature level within the organisation.

Organisations can identify which progression step they have reached in each area and which actions to undertake if they wish to progress on the journey towards sustainable research data. For many organisations, the step 'Developed on organisational level' will be the aspired final destination of their journey.

Organisations that have reached this step and wish to advance even further will find additional guidance under '**Further advancement and alignment**'. This part refers to organisational collaboration with (inter)national partners in order to align approaches and achieve a level playing field (at a national or an international level, with different disciplines).

Definitions of the six areas in the matrices

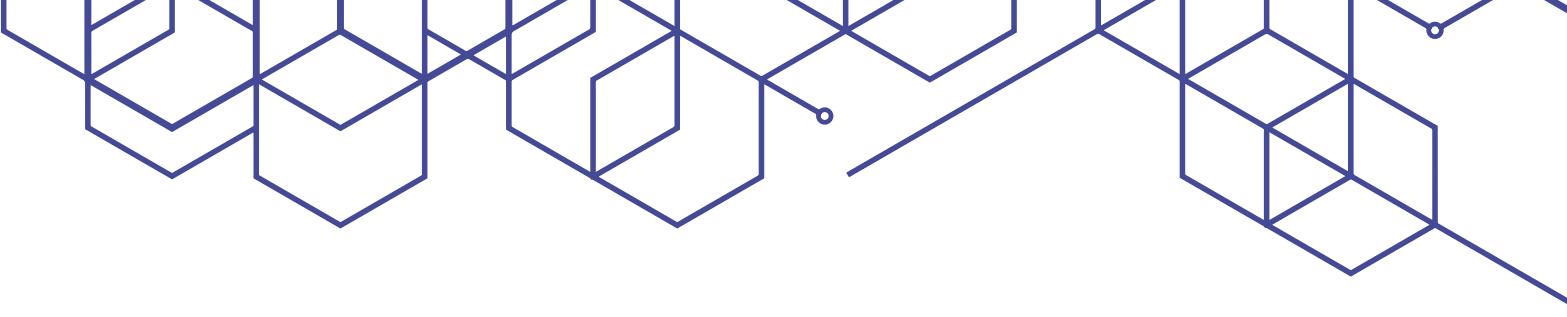
The six areas in which each organisation should take action were defined along the following lines:

ORGANISATIONAL ENGAGEMENT AND COMMITMENT

- For all organisation types, this area refers to the organisation acknowledging the need to develop solutions for sustainable research data and being committed to seek alignment of approaches with other research stakeholders (such as RPOs, RFOs, RDIs, research communities).

POLICY ENVIRONMENT

- For all organisation types, this refers to the organisation clarifying its objectives for data sustainability and interoperability and aiming at coherent policies for all types of organisations.
- For RFOs, this area will cover a range of issues related to RDM that reflect the way of working of the RFO, such as support for RDM infrastructure and/or training. The actual design of the organisational policy will depend on the mission of the RFO and the mandate it has in its national context.
- For RPOs, this area refers to principles and practices on RDM established by the RPO and to be followed by its researchers. The RPO will seek to provide the necessary support to its researchers.
- For RDIs, this area refers to principles and practices on RDM. These include services for researchers and take into account, where needed, disciplinary differences.



FINANCIAL ASPECTS

- For RFOs, this area relates to funding of and investment in RDM and RDIs.
- For RPOs, this area relates to access to funding for the RPO and how the funding is used to support data sharing and interoperability.
- For RDIs, this area refers to the development and implementation of business models for sustainable funding streams.

TRAINING

- For all organisation types: the common understanding of RDM, data sharing, and interoperability is considered a shared responsibility among all organisations. Training comprises both RDM training for researchers and for organisational staff.
- For RFOs, this area relates to the organisation's contributions to building and maintaining skills and competencies for researchers, the organisations it supports, and their own staff involved in RDM.
- For RPOs, this area relates to training and competency enhancement for both researchers and RDM support staff.
- For RDIs, this area relates to training and competency enhancement of RDI staff, to support researchers in the RDM efforts, as well as training for the users of RDIs.

TECHNICAL PREPAREDNESS

- For RFOs, this area relates to investments in the development and implementation of technology to support RDM.
- For RPOs, this area relates to contributions to infrastructures, data hubs, interfaces, and information management issues that ensure interoperability.
- For RDIs, this area relates to professional technical support for data management, including metadata, storage, usage/accessibility, and APIs.

COMMUNICATION AND AWARENESS RAISING

- For RFOs, this area relates to stakeholder engagement and community development, especially with other RFOs, RPOs, and scientific communities.
- For RPOs, this area relates to both researcher engagement as well as engaging with the broader stakeholder community (such as scientific communities, other RPOs) to seek alignment of approaches.
- For RDIs, this area relates to engagement with researchers as users as well as with the researchers' funding organisations and home institutions.



Research Funding Organisations (RFOs):

Maturity matrix for
sustainable research data



Maturity level areas	Progression steps			Further advancement and alignment
	Plans to develop	Development ongoing	Developed on organisational level	
Organisational Engagement and Commitment	<ul style="list-style-type: none"> RFO takes first steps towards considering RDM issues and defining where it needs to act. RFO determines the scope of its activities, including looking to others for guidance and best practices. 	<ul style="list-style-type: none"> RFO sees effective RDM as part of its strategic objectives and develops a comprehensive strategy⁹ to include policy, funding, technical infrastructure, and training as appropriate. RFO ensures that RDM objectives are developed within the context of related organisations. 	<ul style="list-style-type: none"> RFO has an RDM strategy. RFO actively engages in RDM issues, including financial support for policy implementation, training, and, where appropriate, the provision of infrastructure for long-term data preservation. RFO is in dialogue and collaboration with related RFOs, RPOs, and RDIs at (inter)national level to advance practical RDM issues. 	<ul style="list-style-type: none"> RFO seeks alignment on RDM policies and practices amongst RFOs, RPOs, and RDIs at (inter)national level. RFO engages in dialogue and collaboration at (inter)national level for policy, training, provision of RDI and so on. RFO helps to provide a level playing field at (inter)national level.
Policy Environment	<ul style="list-style-type: none"> RFO outlines areas of policy and process for policy development that likely need to be addressed. RFO identifies relevant stakeholders within and outside the RFO in the policy development process. RFO reviews policies from related organisations to ensure consistency (where appropriate and possible). 	<ul style="list-style-type: none"> RFO has an initial RDM policy in place, including guidelines on when and how data should be made available and findable. RFO has established a policy review/revision date, recognising that the policy will need revision and refinement as organisational maturity increases. RFO consults with stakeholders on policy implementation and monitoring. RFO has established dialogue with RPOs and RDIs that collaborate with the RFO to ensure they are developing policies and procedures that reflect the requirements of the RFO. 	<ul style="list-style-type: none"> RFO has a well-developed comprehensive RDM policy and supporting environment, including in research funding peer-review processes. RFO engages with a community of stakeholders for policy implementation, for example, through developing a process to monitor the 'FAIRness' of research data. RFO has a complementary community development process established as part of policy development, including the provision of RDM training and infrastructure.¹⁰ 	<ul style="list-style-type: none"> RFO has aligned its policy with those from international organisations (such as Science Europe, RDA, & OECD), with other (inter)national RFOs, or with a national policy framework for RDM. RFO policy is aligned with evolving legal framework around research data and European Data Strategy. RFO policy is developed in collaboration with RPOs and research communities and in alignment with community norms.¹¹ RFO is part of an international RFO network for continuous exchange on further policy development/alignment.

		<ul style="list-style-type: none"> RFO works towards alignment of policies at (inter)national level through engaging with related organisations. RFO initiates monitoring exercises to measure policy impact. 	<ul style="list-style-type: none"> 'FAIRness' of research data is part of the criteria for research funding evaluations. RFO has monitoring in place to measure policy impact, leading to iterative development of policies. RFO policy includes enhanced guidelines to ensure sustainability.¹²
Financial Aspects	<ul style="list-style-type: none"> RFO acknowledges that it has a role in ensuring that RDM activities, including services, are supported financially¹³ and discusses with RPOs and RDIs what this should entail. RFO acknowledges that RDM is a shared endeavour and that combining resources may lead to economies of scale. RFO makes some funding available but funding is not integrated into a comprehensive RDM strategy. 	<ul style="list-style-type: none"> RFO has defined an initial budget allocation for RDM and areas that should be financed. This might be a short-term, interim response awaiting effective longer-term plans. RFO funds RDM primarily on a project-by-project basis, without an overall investment strategy. Funding might be used to incentivise RDM activities within research projects. RFO actively explores opportunities for co-operation with other stakeholders to realise efficiencies and economies of scale. RFO foresees funding to support training and development of basic and domain-specific RDM skills and awareness raising for all researchers. 	<ul style="list-style-type: none"> RFO acknowledges RDM as a fundable element of the research process. RFO has an enhanced, longer-term investment plan to support its RDM strategy, including the recognition that the RFO must provide funding for the development of training and for RDM infrastructure operations activities beyond the life of a given project or programme. RFO expects that supported RPOs and RDIs strive to gain best value for money through joint programmes, common systems and so on, and collaborates with other RFOs to achieve this. RFO is part of an agreement on an (inter)national investment strategy for RDM to ensure a level playing field between RFOs, RPOs, RDIs, and research communities. RFO provides clear indications on long-term funding periods and allocation of funds. RFO provides funding programmes for training development and activities to sustain the operation of RDM infrastructure. RFO enables/supports publicly funded RDIs to develop sustainable business models, including clarification of the role of the private sector, shared services, and opportunities for commercial use of data.

Training

Note: An RFO will probably not provide training for researchers, however, it will expect RPOs to ensure the researchers they support are effectively trained. In many cases an RFO will provide internal training for their staff.

- RFO acknowledges the importance of RDM training.
- RFO consults with stakeholders (such as RPOs and RDIs) on likely training requirements.
- RFO identifies existing training available.
- RFO clarifies responsibilities for RFO, RPOs, and RDIs with the other organisations and develops training plans for RDM to cross-link to RPO curricula and training undertaken within RPOs and RDIs.
- RFO supports training development on basic and domain-specific RDM skills and awareness raising for all researchers.
- RFO supports the community to develop expected RDM competency levels (in conjunction with RPOs and RDIs). This should include developing a common understanding of the benefits of FAIR data, and compliance with legal requirements (for example in relation to data protection).
- RFO expects training in effective RDM skills as a core element of support provided to researchers. It clarifies responsibilities for training with other organisations and supports the provision of specific advanced training.
- RFO supports the development and implementation of (inter) national RDM competency levels and training curriculum.
- RFO collaborates with RFOs and RPOs on (inter)national level to define responsibilities for training and sources of related funding.

Technical Preparedness

- RFO acknowledges that it has a role in supporting the technical environment for RDM, such as collaborating with other stakeholders (RPOs, RDIs, scientific communities) to identify appropriate standards, and through funding the use of RDIs.
- RFO establishes an internal awareness on management level regarding the allocation of appropriate resources.
- RFO considers the value of joint working shared services between RFOs, RPOs, and RDIs.
- RFO supports work with RPOs, RDIs, and research communities to identify appropriate metadata standards and required infrastructure and services.
- RFO has awareness building measures in place and aims to ensure that disciplinary differences with respect to RDM are recognised in an effective and efficient technical implementation.
- RFO develops mechanisms to support joint working, development of shared services, and so on, to help deliver economies of scale.
- RFO acknowledges that RDM is an (inter)national shared endeavour, best achieved through federation of services, joint working, and joint facilities (such as EOSC).
- RFO has an RDM technical plan in place to ensure FAIR data. The plan includes the provision of RDM infrastructure and the promotion of interoperability as strategic choices.
- RFO agrees with all stakeholders on the limits of responsibility of the RFO, including specific allocation of funds and definition of responsibilities.
- RFO has established communication and co-operation channels with other stakeholders on a national level.
- RFO has established exchanges with other research stakeholders on (inter)national level on technology issues.
- RFO is part of a joined up (inter) national RDM technical approach across all RFOs to ensure RDM infrastructures reflect wider best practice for sustainability.
- RFO aims to provide a level playing field and equal access to RDM infrastructure in collaboration with other organisations, so that access to infrastructure is no longer a barrier to doing effective RDM.
- RFO has strategic planning in place to match RDM requirements and IT infrastructure.

Communication and Awareness Raising

- RFO develops its agenda for research data, in collaboration with other stakeholders (such as government, RPOs).
- RFO plans and initiates a process to engage all stakeholders in the dialogue on the importance of RDM issues.
- RFO develops public statements of intent around the importance of research data and RDM.
- RFO has established communication pathways to engage RPOs, RDIs, and researchers supported by the RFO on the agenda for research data.
- RFO has developed communication materials to show how it addresses the issue strategically and to underline that RDM is a shared endeavour by all stakeholders.
- RFO has identified and highlighted best practices/projects from its funding activities and provides information about associated costs.
- RFO collaborates with other related organisations to help develop and advance the agenda for research data.
- RFO acknowledges that RDM is part of research processes and should not be seen as anything out of the ordinary.
- Collaboration, shared services, and so on become the default. A default funding model should enhance existing provisions rather than developing new ones, unless there is a very clear business case for this.
- RFO provides well-developed communication materials and pathways, stressing the importance of RDM, the roles of each key stakeholder, and the commitment of the RFO towards RDM (the agenda for research data).
- RFO clearly communicates the support RFOs offer for RDM capacity building and specialist careers development.
- RFO actively engages with the community to make RDM a shared endeavour across the RFO's research community.
- RFO has established effective RDM as part of research processes and does not see it as something out of the ordinary.
- RFO uses the agenda for research data to communicate objectives to own researchers and staff and beyond.
- RFO develops innovative business models involving private sector participation.





Maturity level areas	Progression steps			Further advancement and alignment
	Plans to develop	Development ongoing	Developed on organisational level	
Organisational Engagement and Commitment	<ul style="list-style-type: none"> RPO is engaged with RDM, at least to the level of statements of intent on the importance of research data and RDM. RPO determines the scope of its activities and is in contact with its stakeholders to acquire high-level recommendations/guidelines and possibly looking to others for guidance and best practices. 	<ul style="list-style-type: none"> RPO is actively engaged with RDM issues and has defined long-term handling of research data and related outputs as a priority. RPO commits to addressing RDM issues, through the development of a comprehensive strategy, to include policy, funding, technical infrastructure, and training. 	<ul style="list-style-type: none"> RPO is effectively engaged with RDM issues and has a strategy for RDM in place. RPO has active support in place, including financial support, support for policy implementation, technical support (for examples via the availability of data stewards), training, and where appropriate the provision of long-term data infrastructure. RPO has established a stable communication channel with technical providers (at least one between RDIs or EU/country wide services). RPO is in dialogue and collaboration with related RFOs, RPOs, and RDIs at (inter)national level to advance on practical RDM issues. 	<ul style="list-style-type: none"> RPO enhances engagement on RDM at (inter)national level; the policies are fully aligned with RFOs, RDIs, scientific communities, and government policies. RPO has connections with technical providers (be they RDIs, country-wide services) that include a long-term commitment and are shared with similar RPOs.
Policy Environment	<ul style="list-style-type: none"> RPO identifies areas for RDM policy development, establishes respective processes, and consults with other stakeholders (RFOs, RDIs, government) 	<ul style="list-style-type: none"> RPO has an initial RDM policy in place. RPO has a policy review/revision date established, recognising that policy will need revision and refinement as organisational maturity increases. 	<ul style="list-style-type: none"> RPO has well-developed, comprehensive policies within a supporting environment (data processing procedures throughout and after finalisation of the research process, compliance monitoring, including as part of the research funding peer-review process at project approval). 	<ul style="list-style-type: none"> RPO's RDM policies and guidelines for the utilisation of selected RDIs (if outsourced) are fully aligned with other (inter)national RPOs and with stakeholders (RFOs, RDIs, scientific communities, Government).

	<ul style="list-style-type: none"> RPO consults with stakeholders on policy implementation and monitoring. RPO has established a basic support for researchers, in co-ordination with RFOs, to meet funder and legal requirements. RPO considers RDM requirements for future assessment processes of internal projects. RPO negotiates with selected RDIs to handle the research outputs (if outsourcing has been decided); otherwise, an internal process of procurement and deployment is planned. 	<ul style="list-style-type: none"> RPO offers selected RDIs (in-house, nationwide, EU-wide) to researchers to handle the research outputs; MoUs or similar are available with these RDIs. In case an in-house solution is selected, it is declared operational and usable. RPO provides clear instructions for support to research and technical staff. RPO has a complementary community development process established as part of policy development, including the provision of RDM training and infrastructure.¹⁰ 	<ul style="list-style-type: none"> RPO ensures alignment of its policies with evolving legal framework around research data and European Data Strategy. RPO has set up a co-ordination process with peer organisations. RPO has set up effective long-term management agreements between RDIs and RFOs. RPO collaborates with disciplines to seek broader alignment.
--	--	--	--

Financial Aspects

- | | | | |
|--|--|---|--|
| <ul style="list-style-type: none"> RPO acknowledges that it has a role in ensuring RDM activities/infrastructure are supported financially. RPO discusses with RFOs, government, and RDIs (as service providers) to allocate appropriate levels of funding, including exploring potential new sources. RPO has initiated discussions with research communities on consequences for project evaluation modalities, including discussions on the eligibility of research costs. | <ul style="list-style-type: none"> RPO has defined an initial budget allocation for RDM and areas for spending. RPO has allocated funding, internally and with other entities (RFOs, government, RDIs). This might be a short-term interim response, awaiting a consolidated longer-term plan. | <ul style="list-style-type: none"> RPO has an enhanced, longer-term investment plan to support research communities with implementing RDM strategies. This includes the recognition that RPO needs to provide training and first-level RDM infrastructure beyond the life of any specific project or programme. RPO needs to include potentially modified evaluation modalities for projects including incentives for researchers' career paths. RPO has a process to monitor efforts and investments, and assesses and solves gaps. | <ul style="list-style-type: none"> RPO foresees budgets to meet RDM policies and plans for most research groups. RPO foresees budget also for project evaluation (at approval, throughout the research project, and after finalisation of the project) and for the impact of researchers' career paths. RPO secures additional funding with RFOs/government or carves it out of the standard budget. RPO has a support service that is sustainable in the long-term. |
|--|--|---|--|

Training

- RPO acknowledges the importance of RDM training and consults communities on training requirements and RDIs on training needs.
- RPO provides a basic training environment for researchers and support staff, usually via internal expertise in the RPO.
- RPO has initiated evaluating the possibility of engaging data stewards and preparing an internal consultancy service.
- RPO provides a specific support service for researchers, via data stewards and librarians, available to help research groups in all project phases.
- RPO has defined expected RDM core competency levels for the research community and established them throughout the organisation.
- RPO considers continuous/regular training in effective RDM skills (internal or served via larger communities such as RFOs or science communities) a core element of support for researchers and a recurrent activity in the long-term planning.
- RPO has internal or outsourced experts (such as data stewards, digital curation experts, librarians, and archivists) available for consultancy in all phases of the research process and beyond to advise on RDM planning, data curation, and storage.
- RPO's training activities cover legal aspects, (inter)national standards for technical aspects, RDM policies, life-cycle data management, and digital preservation.
- RPO has a complete training and support system with sufficient staff, with specific support personnel offered to research groups (data stewards, librarians, and so on).
- RPO's training activities cover legal aspects, (inter)national standards for technical aspects, RDM policies, life-cycle data management, and digital preservation.

Technical Preparedness

- RPO acknowledges that it has a role in technical support for RDM, for example, on metadata standards or provision of RDIs, and collaborates with other stakeholders (RFOs, RDIs, scientific communities) to identify appropriate standards.
- RPO investigates possible RDIs to be used as technical solution to handle research data, including links between own infrastructure and network with other RDIs.
- RPO works with its research communities to identify appropriate metadata standards and required infrastructure (internally/externally/networking) addressing the respective disciplinary needs.
- RPO conducts tests with selected RDIs to handle the research outputs, including permanent feedback with its own researchers.
- RPO decides whether to outsource the technical aspect or keep it in-house.
- RPO has an RDM technical plan to ensure FAIR data. It includes metadata standards and provision of RDM infrastructure in collaboration with technical service providers. Responsibilities within the organisation (clear governance structure) and budget allocation are crucial elements of the plan.
- RPO considers FAIR handling of research data the norm.
- RPO is part of a joined-up RDM technical approach across all RPOs in the same field/country.
- RPO ensures, in case an RPO-owned infrastructure is chosen, that it provides adequate support staff, is fully operational, and/or ensures seamless data flow to national/international infrastructure providing long-term handling of research data.
- RPO is part of an agreed and established governance of research data networks.

Communication and Awareness Raising

- RPO develops its agenda for research data, in collaboration with the stakeholders (such as RFOs, government).
- RPO sets up a communication process to engage own researchers on the value of FAIR research data and RDM.
- RPO has established communication pathways to engage research groups it supports; aspects like awards and better career opportunities are clarified.
- RPO provides information showing its commitment and supporting role for researchers.
- RPO provides well-developed communication materials and pathways, stressing the importance of RDM, the roles of each key stakeholder, and the commitment of the RPO towards RDM.
- RPO's active community engagement leads to RDM becoming a shared and recognised value across the RPO's research communities.
- RPO clearly communicates the objective of FAIRness in its strategy.
- RPO has entered a co-ordination process with other RPOs (nationally and internationally) to develop a common agenda for research data. This needs to be aligned with RFOs' policies and with RDIs' technical offerings.
- RPO has established RDM as part of the everyday activity of research. Effective RDM becomes the 'new normal'.



Research Data Infrastructures (RDI):

Maturity matrices for
sustainable research data



Maturity level areas	Progression steps			Further advancement and alignment
	Plans to develop	Development ongoing	Developed on organisational level	
Organisational Engagement and Commitment	<ul style="list-style-type: none"> RDI establishes links and collaborations and engages with RPOs, RFOs, and/or other RDIs to develop the portfolio of services called for/needed. 	<ul style="list-style-type: none"> RDI has established links and collaborations with RPOs, RFOs, and/or other RDIs (and government where applicable) and defines the services the RDI could provide. 	<ul style="list-style-type: none"> RDI starts to engage on RDM at (inter)national level. RDI provides a portfolio of services ready to be deployed as called for/needed by RPOs and/or RFOs. 	<ul style="list-style-type: none"> RDI aligns on RDM at (inter)national level through collaboration between RPOs, RFOs, RDIs, and research communities. RDI provides co-ordinated procedures with other RDI to develop long-term sustainability and exit scenarios, for example through seamless exchange of data holdings. RDI is connected to, and where applicable co-operating with, other relevant (inter)national RDM initiatives.
Policy Environment	<ul style="list-style-type: none"> RDI has plans to develop policy for sustainable data management. Policies include measures to ensure FAIR data and continuity of data access and services. RDI has started to collaborate with relevant actors (RFO, RPOs, RDIs) in policy development. 	<ul style="list-style-type: none"> RDI policy is developed in consultation with RPOs and/or RFOs where relevant. RDI collaborates with relevant RDIs to improve the alignment of policies and technology. RDI develops agreements with other RDIs for long-term continuity of data access and services. 	<ul style="list-style-type: none"> RDI has fully implemented policies. RDI has signed MoUs or similar with RPOs where relevant. 	<ul style="list-style-type: none"> RDI has policies aligned with those of related RPOs and/or RFOs and other RDIs, as well as in a wider policy framework (such as EOSC). RDI's policies are an integrated part of the total (inter)national RDM policy environment. RDI collaborates with research communities/stakeholders with well-developed links for feedback to improve policies.

Financial Aspects

- RDI seeks partnering and collaboration with RPOs, RFOs and/or government to offer solutions for the sustainability of research outputs when relevant.
- RDI has an understanding of costs and funding mechanism and has tested these assumptions.
- RDI has initiated dialogue with relevant stakeholders for clarification of accountabilities.
- RDI has a regular funding stream, built from short-term projects and one-off funds.
- RDI negotiates with potential partners/collaborators, RFOs, and RPOs on building critical mass funding.
- RDI has started to develop a business model for long-term funding.
- RDI's risk assessment allows funding beyond the 'event horizon' of data loss and RDI identifies the timeframe of acceptable and unacceptable data risks.
- RDI has a sustainable short-term source of financing established.
- RDI explores further business models for raising financial contributions.
- RDI works to have complementary services to other relevant RDIs to make the RDI ecosystem financially robust.
- RDI has plans for closure including a secure budget for data inventory transfer to other repositories if the funding stream fails.
- RDI has a sustainable source of both long- and short-term financing.
- RDI has complementary services to other RDIs to ensure cost efficiency.
- RDI is federated to EOSC or other relevant organisations to help improve access to funding.

Training

- RDI plans to establish links/co-operation with RPOs and/or RFOs and start discussions about the solutions the RPOs and/or RFOs need from the RDI.
- RDI establishes links and engaging/collaborating with RFOs and RPOs to understand support level and training needs.
- RDI prepares a draft portfolio, including for example specialist, technical, and IT skills needed for FAIR data handling.
- RDI takes measures to assure, in collaboration with RPOs, that the users of the RDI's services can comply with the RDI's demands for the quality of research data, methodologies, and processes.
- RDI has aligned to relevant RFOs and/or RPOs requirements and/or policy nationally.
- RDI provides user guides and support information (helpdesk).
- RDI offers easily accessible training courses.
- RDI ensures sufficiently trained staff is available.
- RDI has training requirements aligned with (inter)national policies.
- RDI provides a portfolio of solutions (such as training, documentation).
- RDI offers specialist training activities (for example, on FAIR data, technical and IT skills, and financial and legal skills for data management services).

Technical Preparedness

- RDI explores technical solutions for alignment with RPOs, relevant scientific communities, collaborators, and/or other RDIs.
- RDI ensures alignment with technology developments supporting the wider use of scientific data.
- RDI considers support for FAIR data the norm.
- RDI provides data management, networking, and storage services.
- RDI is an active member of a national RDI network, exchanging experience on technology and aiming to ensure best practice for sustainability.

Communication and Awareness Raising

- RDI starts to plan outreach to other stakeholders.
 - RDI starts to engage with the user communities.
 - RDI has relevant technology in place, but only on a low to medium level, and basic functionalities.
 - RDI acknowledges the need for user-/discipline-specific requirements on technology and services.
 - RDI plans to collaborate with other stakeholders/users to establish development priorities.
 - RDI ensures findability and accessibility of datasets.
 - RDI has strategies to ensure that the technological solutions are maintained to relevant industry standards.
 - RDI has developed plans to implement interoperability standards and implementation has started.
 - RDI has established and/or is participating in a co-operative (inter)national RDI network.
 - RDI has still some potential for further optimisation with respect to disciplinary metadata and standards.
 - RDI has an exit plan for ensuring sustainable access to research data (transfer to other repositories).
 - RDI has implemented technology using the highest quality standards, including frequent updates and further professionalisation perspectives.
 - RDI has a mid- to long-term strategic technological planning to ensure stable operability and service.
 - RDI participates in (inter)national consortia or organisations of computational centres where relevant.
 - RDI actively develops and applies modern, strategic technologies (AI, blockchain, and so on) for better data management and re-use possibilities, possibly in collaboration with other RDIs.
 - RDI provides co-ordinated procedures with other RDIs to swap data holdings potentially seamlessly if necessary.
 - RDI provides targeted services (such as data curation, long-term preservation, data storage, computing facilities) to research groups or institutions.
 - RDI has complementary services to other RDIs.
- RDI develops a communications plan on which stakeholders they aim to reach and how.
 - RDI develops a communications plan to address user communities, RFOs, and RPOs.
 - RDI communicates available RDI services and access policies and requirements for the quality of research data, methodologies, and processes.
 - RDI has a consolidated communications plan available to the user communities, RFOs, and RPOs.
 - RDI has identified the relevant stakeholders and defined communication channels with these stakeholders.
 - RDI uses appropriate communications channels to present examples of successful use of research data (best practice examples).



References

- 1 Data in this context are defined not only as the final outcomes of a project, but also corresponding aspects, such as the software needed, raw data, and metadata.
- 2 Data should be Findable, Accessible, Interoperable, and Re-usable.
- 3 The TRUST principles for digital repositories refer to Transparency, Responsibility, User focus, Sustainability, and Technology as guiding principles that demonstrate the trustworthiness of digital repositories. Especially relevant in the context of this guide is the principle of Sustainability, defined as sustaining services and preserving data holding for the long-term.
- 4 RDIs include data-related research infrastructures and elnfrastuctures.
- 5 External organisations present at the validation workshop were CERN, CESSDA, Charles University Prague, CLARINO – Common Language Resource and technology Infrastructure, CSC – IT Centre for Finland, DARIAH-EU, Digital Curation Centre, Digital Preservation Coalition, ELLETTRA, EUA – The European University Association, Friedrich-Schiller University Jena, Medical University of Graz, Norwegian Meteorological Institute, Swiss Institute of Bioinformatics, The Guild of European Research-Intensive Universities, TU Delft, University of Coimbra, University of Edinburg, University of Geneva, University of Pavia, University of Turku, University of Turku Library, Uppsala University, Vilnius University, Vytautas Magnus University, Wellcome, and 4TU.ResearchData.
- 6 <https://scienceeurope.org/our-priorities/research-data/>
- 7 A European Strategy for data (COM(2020) 66 final): <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1593073685620&uri=CELEX%3A52020DC0066>
- 8 The strategic goals can refer to the timing of taking future actions as well as the decision of which progression step an organisation wants to reach.
- 9 The RDM strategy is a core element that cuts across many of the thematic areas. It brings together separate thematic activities into a more organised whole and should not be developed in isolation.
- 10 Implementation of policy must go hand-in-hand with establishment of necessary skills and infrastructure.
- 11 The aim is to provide a level playing field and consistent policy environment for all researchers regardless of funder or discipline.
- 12 Including issues of interoperability and re-usability of data, necessary for making informed judgements of what data are likely to have a long-term value, and identification of roles and responsibilities.
- 13 For example, RFO provides RDM funding in project budget and/or includes funding in its data sharing policies.



Science Europe
Rue de la Science 14
1040 Brussels
Belgium

Tel +32 (0)2 226 03 00
Fax +32 (0)2 226 03 01
office@scienceeurope.org
www.scienceeurope.org