UNIVERSITY OF AMSTERDAM

Promoting Open Access

Frans Oort, University of Amsterdam

EOSC National Tripartite Event Netherlands, 22 May 2024, Utrecht

Contact address: f.j.oort@uva.nl

Open Access Policy

Green OA
Diamond OA
Gold OA

Self-archiving in compliance with publisher's policy

Full, immediate OA publishing, without Article Processing Charge - APC

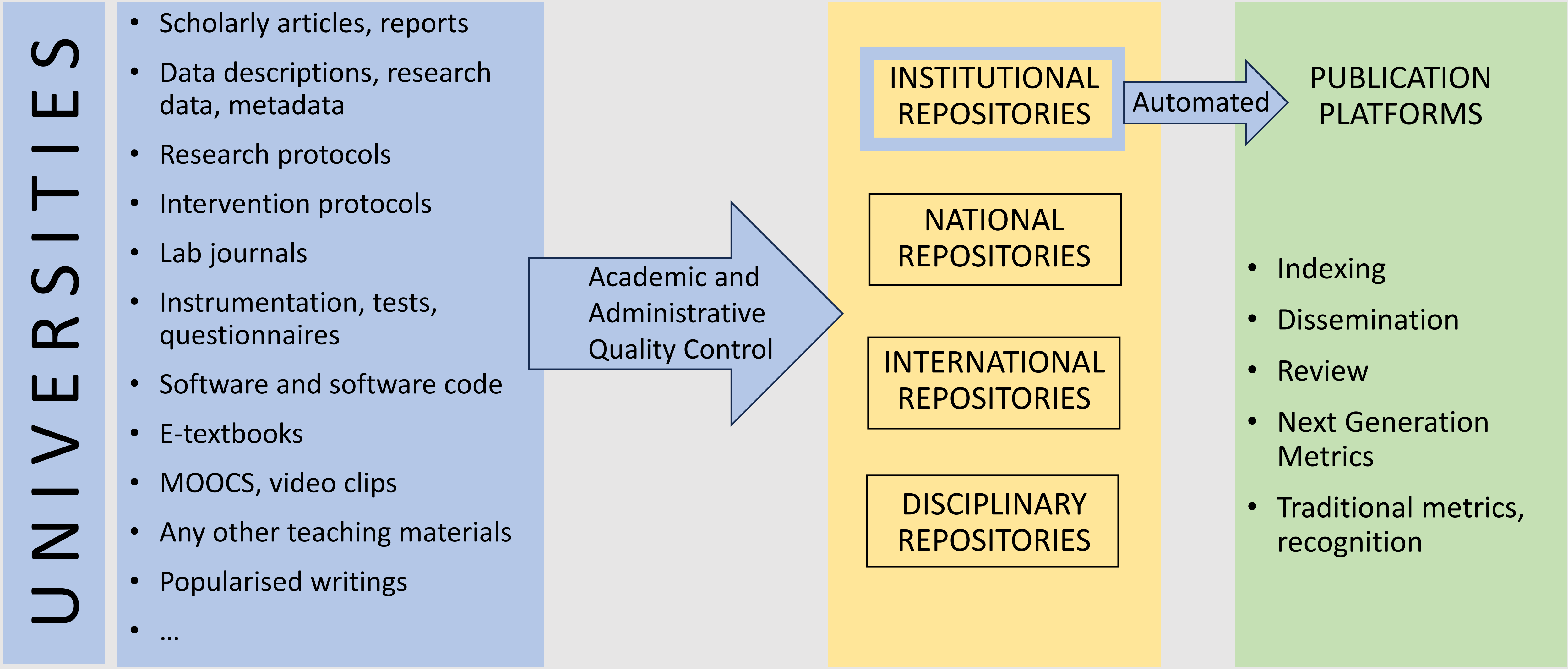
Full, immediate OA. Author pays an APC

- **Discourage Gold OA**
  except through general Read & Publish agreements
- **Discourage Hybrid OA**
  prohibiting the use of first flow funding
- **Facilitate Green OA**
  using university repository (justified by Dutch Copyright Act)
- **Promote Diamond OA**
  through various funds
- **Advocate Public Infrastructure**
  starting projects, influencing policies, engaging organisations

Components of public infrastructure

- Repositories
- Publication platforms
- Research Data Exchange (RDX)
- Open Knowledge Base (OKB)
- Research Information Systems (CRIS)

All academic output can be published in repositories



UNIVERSITIES

- Scholarly articles, reports
- Data descriptions, research data, metadata
- Research protocols
- Intervention protocols
- Lab journals
- Instrumentation, tests, questionnaires
- Software and software code
- E-textbooks
- MOOCS, video clips
- Any other teaching materials
- Popularised writings
- …

Academic and Administrative Quality Control

INSTITUTIONAL REPOSITORIES

NATIONAL REPOSITORIES

INTERNATIONAL REPOSITORIES

DISCIPLINARY REPOSITORIES

Automated

PUBLICATION PLATFORMS

- Indexing
- Dissemination
- Review
- Next Generation Metrics
- Traditional metrics, recognition

University Journals — – A Publication Platform for All Scholarly Output

About University Journals **For Universities** **For Researchers** **Contact us**

University Journals

LIBER QUARTERLY
The Journal of the Association of European Research Libraries
Vol. 30 (2020) 1–15 | e-ISSN: 2213-056X

UNIVERSITY JOURNALS.

Consolidating Institutional Repositories in a Digital, Free, Open Access Publication Platform for *All* Scholarly Output

Saskia Woutersen-Windhouwer

University Library, University of Amsterdam/Leiden University Libraries, Leiden University, The Netherlands
s.woutersen@library.leidenuniv.nl, orcid.org/0000-0003-0120-266X

Eva Méndez Rodríguez

Library and Information Sciences Department, Universidad Carlos III de Madrid, Spain
emendez@bib.uc3m.es, orcid.org/0000-0002-5337-4722

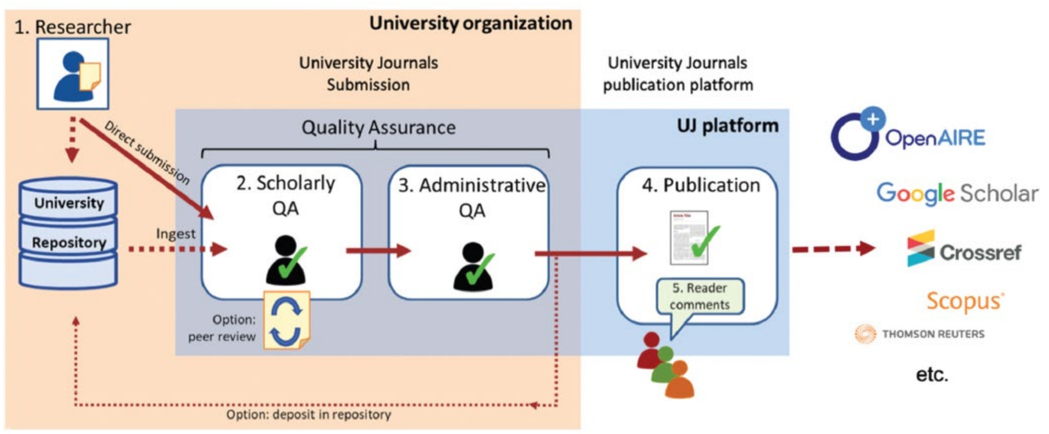
Jeroen Sondervan

Utrecht University Library, Utrecht University, The Netherlands
j.sondervan@uu.nl, orcid.org/0000-0002-9866-0239

Frans J. Oort

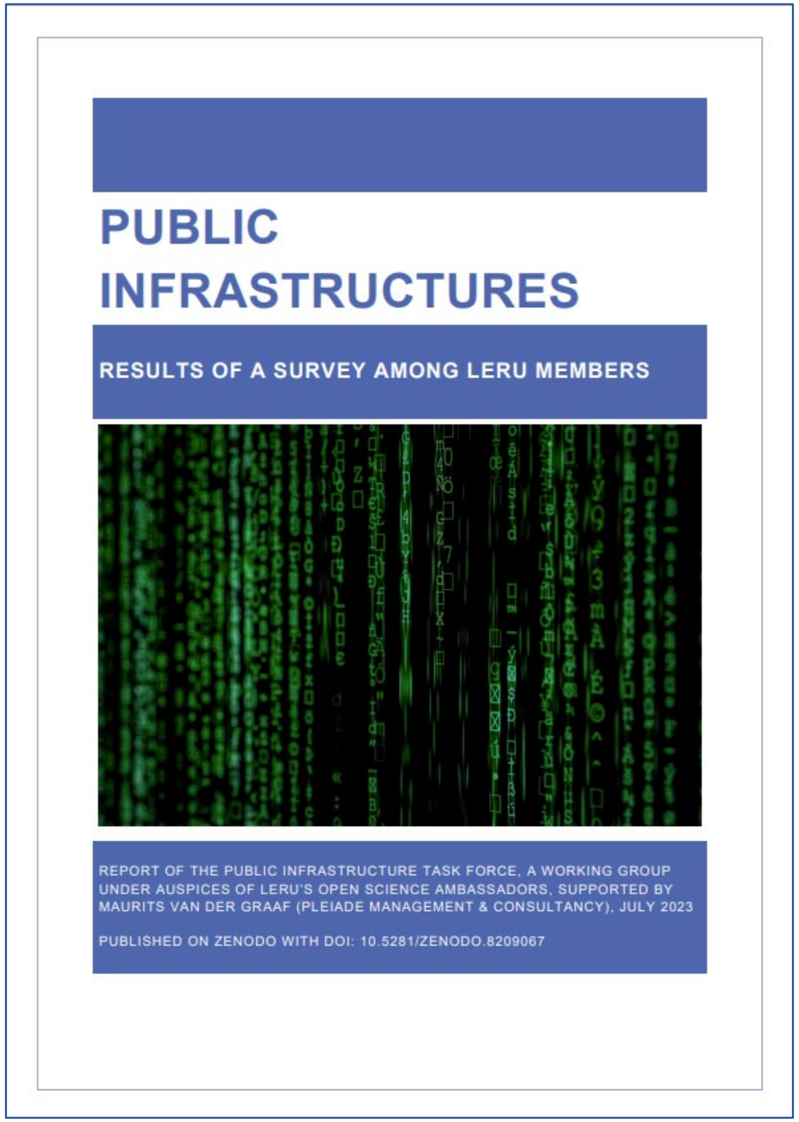
Research Advisory Committee, University of Amsterdam, The Netherlands
f.j.oort@uva.nl, orcid.org/0000-0003-1823-7105

Fig. 1: Workflow University Journals (by Max Haring).



LERU Survey (www.leru.org)

On the Need to Establish Public Infrastructure to Preserve Digital Sovereignty

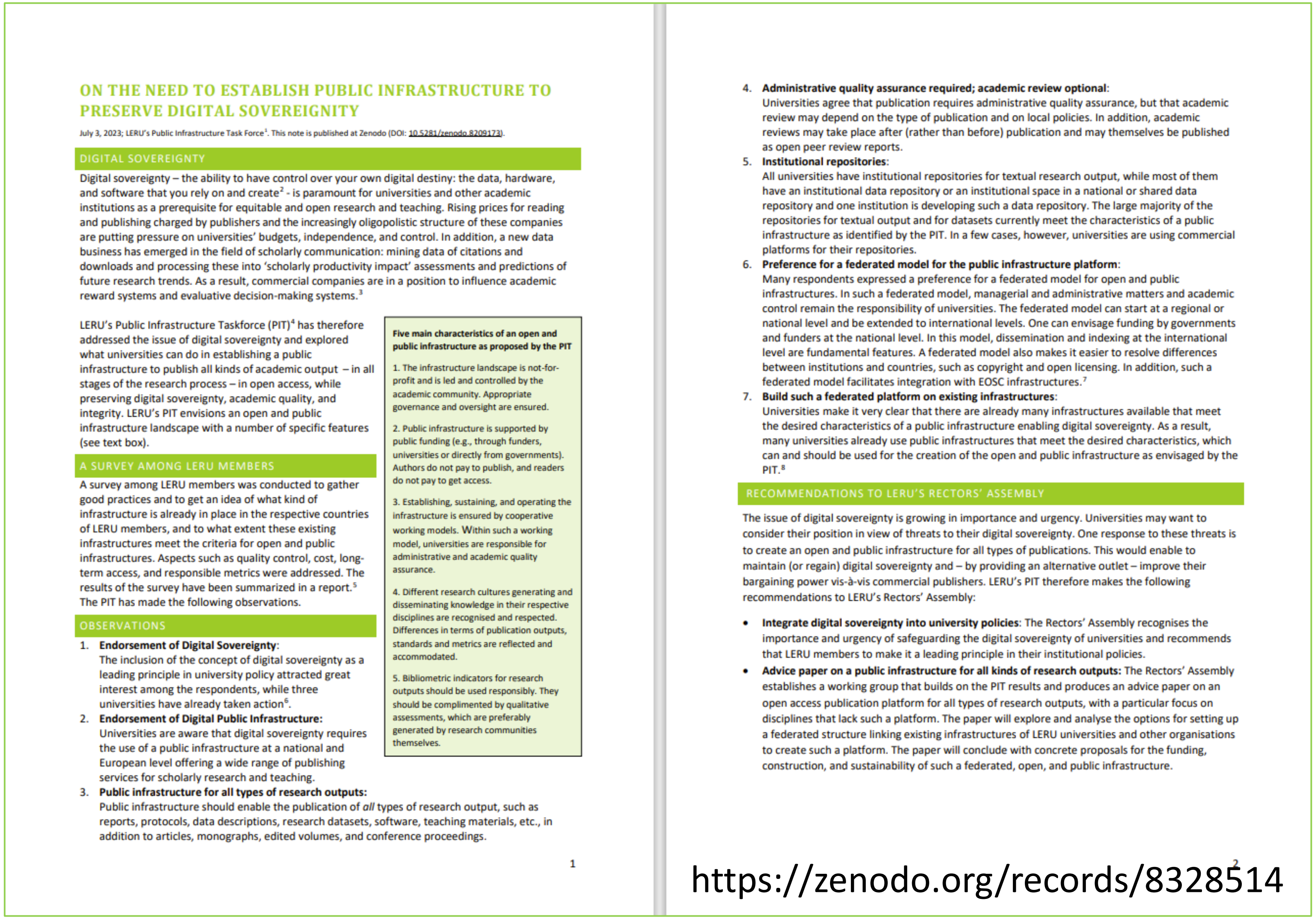


PUBLIC INFRASTRUCTURES

RESULTS OF A SURVEY AMONG LERU MEMBERS

REPORT OF THE PUBLIC INFRASTRUCTURE TASK FORCE, A WORKING GROUP UNDER AUSPICES OF LERU'S OPEN SCIENCE AMBASSADORS, SUPPORTED BY MAURITS VAN DER GRAAF (PLEIADE MANAGEMENT & CONSULTANCY), JULY 2023

PUBLISHED ON ZENODO WITH DOI: 10.5281/ZENODO.8209067



ON THE NEED TO ESTABLISH PUBLIC INFRASTRUCTURE TO PRESERVE DIGITAL SOVEREIGNITY

July 3, 2023; LERU's Public Infrastructure Task Force[1]. This note is published at Zenodo (DOI: 10.5281/zenodo.8209173).

DIGITAL SOVEREIGNTY

Digital sovereignty – the ability to have control over your own digital destiny: the data, hardware, and software that you rely on and create[2] - is paramount for universities and other academic institutions as a prerequisite for equitable and open research and teaching. Rising prices for reading and publishing charged by publishers and the increasingly oligopolistic structure of these companies are putting pressure on universities' budgets, independence, and control. In addition, a new data business has emerged in the field of scholarly communication: mining data of citations and downloads and processing these into 'scholarly productivity impact' assessments and predictions of future research trends. As a result, commercial companies are in a position to influence academic reward systems and evaluative decision-making systems.[3]

LERU's Public Infrastructure Taskforce (PIT)[4] has therefore addressed the issue of digital sovereignty and explored what universities can do in establishing a public infrastructure to publish all kinds of academic output – in all stages of the research process – in open access, while preserving digital sovereignty, academic quality, and integrity. LERU's PIT envisions an open and public infrastructure landscape with a number of specific features (see text box).

Five main characteristics of an open and public infrastructure as proposed by the PIT

1. The infrastructure landscape is not-for-profit and is led and controlled by the academic community. Appropriate governance and oversight are ensured.

2. Public infrastructure is supported by public funding (e.g., through funders, universities or directly from governments). Authors do not pay to publish, and readers do not pay to get access.

3. Establishing, sustaining, and operating the infrastructure is ensured by cooperative working models. Within such a working model, universities are responsible for administrative and academic quality assurance.

4. Different research cultures generating and disseminating knowledge in their respective disciplines are recognised and respected. Differences in terms of publication outputs, standards and metrics are reflected and accommodated.

5. Bibliometric indicators for research outputs should be used responsibly. They should be complimented by qualitative assessments, which are preferably generated by research communities themselves.

A SURVEY AMONG LERU MEMBERS

A survey among LERU members was conducted to gather good practices and to get an idea of what kind of infrastructure is already in place in the respective countries of LERU members, and to what extent these existing infrastructures meet the criteria for open and public infrastructures. Aspects such as quality control, cost, long-term access, and responsible metrics were addressed. The results of the survey have been summarized in a report.[5] The PIT has made the following observations.

OBSERVATIONS

1. **Endorsement of Digital Sovereignty**:
   The inclusion of the concept of digital sovereignty as a leading principle in university policy attracted great interest among the respondents, while three universities have already taken action[6].
2. **Endorsement of Digital Public Infrastructure:**
   Universities are aware that digital sovereignty requires the use of a public infrastructure at a national and European level offering a wide range of publishing services for scholarly research and teaching.
3. **Public infrastructure for all types of research outputs:**
   Public infrastructure should enable the publication of *all* types of research output, such as reports, protocols, data descriptions, research datasets, software, teaching materials, etc., in addition to articles, monographs, edited volumes, and conference proceedings.

4. **Administrative quality assurance required; academic review optional**:
   Universities agree that publication requires administrative quality assurance, but that academic review may depend on the type of publication and on local policies. In addition, academic reviews may take place after (rather than before) publication and may themselves be published as open peer review reports.
5. **Institutional repositories**:
   All universities have institutional repositories for textual research output, while most of them have an institutional data repository or an institutional space in a national or shared data repository and one institution is developing such a data repository. The large majority of the repositories for textual output and for datasets currently meet the characteristics of a public infrastructure as identified by the PIT. In a few cases, however, universities are using commercial platforms for their repositories.
6. **Preference for a federated model for the public infrastructure platform**:
   Many respondents expressed a preference for a federated model for open and public infrastructures. In such a federated model, managerial and administrative matters and academic control remain the responsibility of universities. The federated model can start at a regional or national level and be extended to international levels. One can envisage funding by governments and funders at the national level. In this model, dissemination and indexing at the international level are fundamental features. A federated model also makes it easier to resolve differences between institutions and countries, such as copyright and open licensing. In addition, such a federated model facilitates integration with EOSC infrastructures.[7]
7. **Build such a federated platform on existing infrastructures**:
   Universities make it very clear that there are already many infrastructures available that meet the desired characteristics of a public infrastructure enabling digital sovereignty. As a result, many universities already use public infrastructures that meet the desired characteristics, which can and should be used for the creation of the open and public infrastructure as envisaged by the PIT.[8]

RECOMMENDATIONS TO LERU'S RECTORS' ASSEMBLY

The issue of digital sovereignty is growing in importance and urgency. Universities may want to consider their position in view of threats to their digital sovereignty. One response to these threats is to create an open and public infrastructure for all types of publications. This would enable to maintain (or regain) digital sovereignty and – by providing an alternative outlet – improve their bargaining power vis-à-vis commercial publishers. LERU's PIT therefore makes the following recommendations to LERU's Rectors' Assembly:

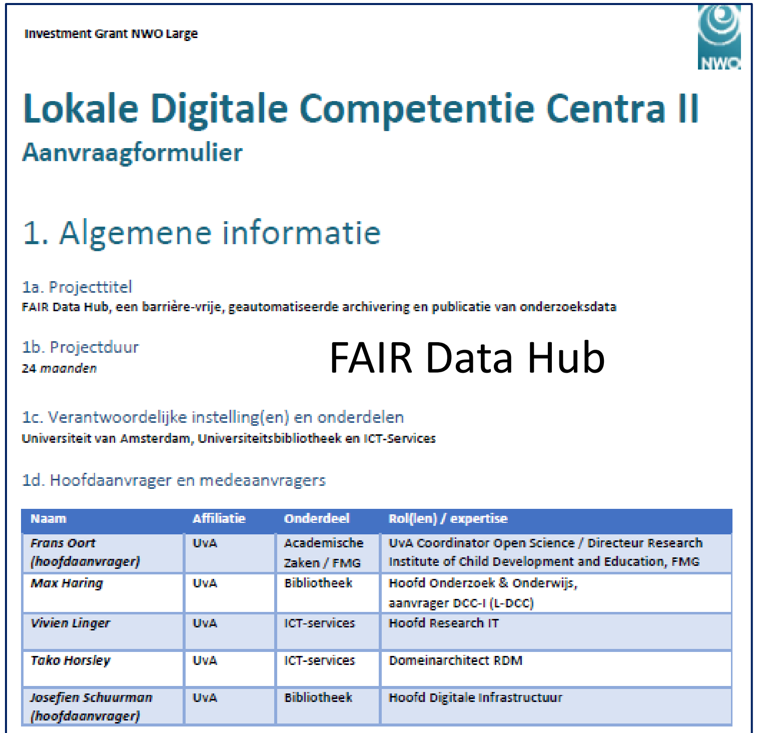
- **Integrate digital sovereignty into university policies**: The Rectors' Assembly recognises the importance and urgency of safeguarding the digital sovereignty of universities and recommends that LERU members to make it a leading principle in their institutional policies.
- **Advice paper on a public infrastructure for all kinds of research outputs:** The Rectors' Assembly establishes a working group that builds on the PIT results and produces an advice paper on an open access publication platform for all types of research outputs, with a particular focus on disciplines that lack such a platform. The paper will explore and analyse the options for setting up a federated structure linking existing infrastructures of LERU universities and other organisations to create such a platform. The paper will conclude with concrete proposals for the funding, construction, and sustainability of such a federated, open, and public infrastructure.

https://zenodo.org/records/8328514

National Coordination Point Research Data Management

Data sovereignty, data governance and digital sovereignty

LCRDM



Investment Grant NWO Large

Lokale Digitale Competentie Centra II

Aanvraagformulier

1. Algemene informatie

1a. Projecttitel

FAIR Data Hub, een barrière-vrije, geautomatiseerde archivering en publicatie van onderzoeksdata

1b. Projectduur

24 maanden

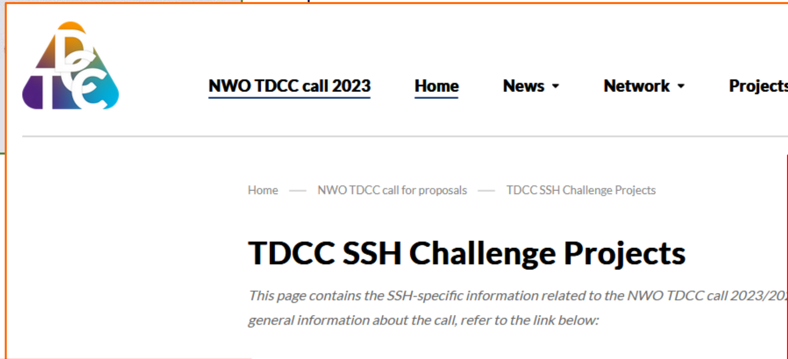
FAIR Data Hub

1c. Verantwoordelijke instelling(en) en onderdelen

Universiteit van Amsterdam, Universiteitsbibliotheek en ICT-Services

1d. Hoofdaanvrager en medeaanvragers

Naam	Affiliatie	Onderdeel	Rol(len) / expertise
Frans Oort (hoofdaanvrager)	UvA	Academische Zaken / FMG	UvA Coordinator Open Science / Directeur Research Institute of Child Development and Education, FMG
Max Haring	UvA	Bibliotheek	Hoofd Onderzoek & Onderwijs, aanvrager DCC-I (L-DCC)
Vivien Linger	UvA	ICT-services	Hoofd Research IT
Tako Horsley	UvA	ICT-services	Domeinarchitect RDM
Josefien Schuurman (hoofdaanvrager)	UvA	Bibliotheek	Hoofd Digitale Infrastructuur



NWO TDCC call 2023 Home News Network Projects

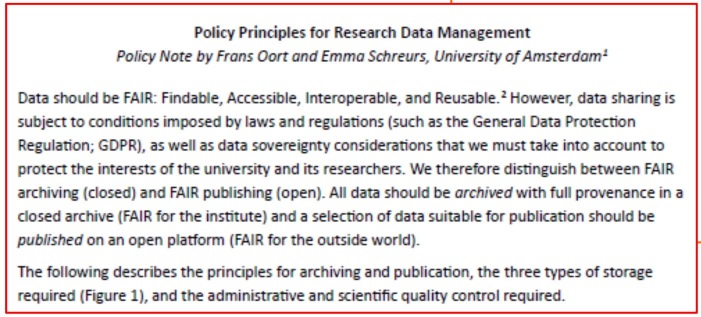
Home — NWO TDCC call for proposals — TDCC SSH Challenge Projects

TDCC SSH Challenge Projects

This page contains the SSH-specific information related to the NWO TDCC call 2023/20
general information about the call, refer to the link below:

General information about the NWO TDCC call

Develop idea & find partners
Submit project idea to TDCC
TDCC-SSH
Submission of final proposal to NWO
NWO
Before 24 May
Before 14 Nov

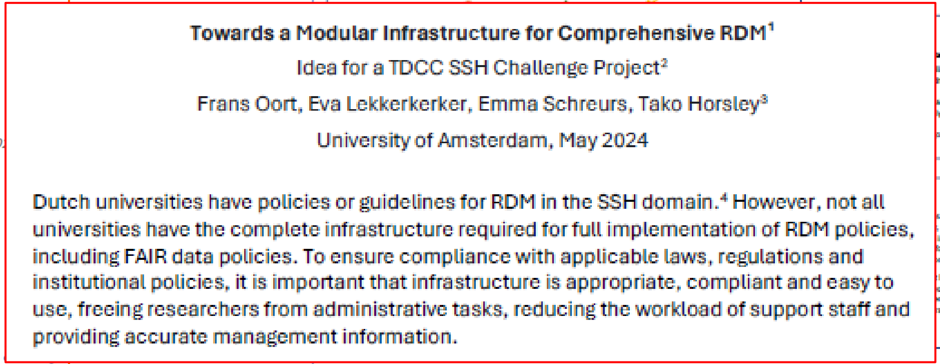


Policy Principles for Research Data Management

Policy Note by Frans Oort and Emma Schreurs, University of Amsterdam[1]

Data should be FAIR: Findable, Accessible, Interoperable, and Reusable.[2] However, data sharing is subject to conditions imposed by laws and regulations (such as the General Data Protection Regulation; GDPR), as well as data sovereignty considerations that we must take into account to protect the interests of the university and its researchers. We therefore distinguish between FAIR archiving (closed) and FAIR publishing (open). All data should be *archived* with full provenance in a closed archive (FAIR for the institute) and a selection of data suitable for publication should be *published* on an open platform (FAIR for the outside world).

The following describes the principles for archiving and publication, the three types of storage required (Figure 1), and the administrative and scientific quality control required.



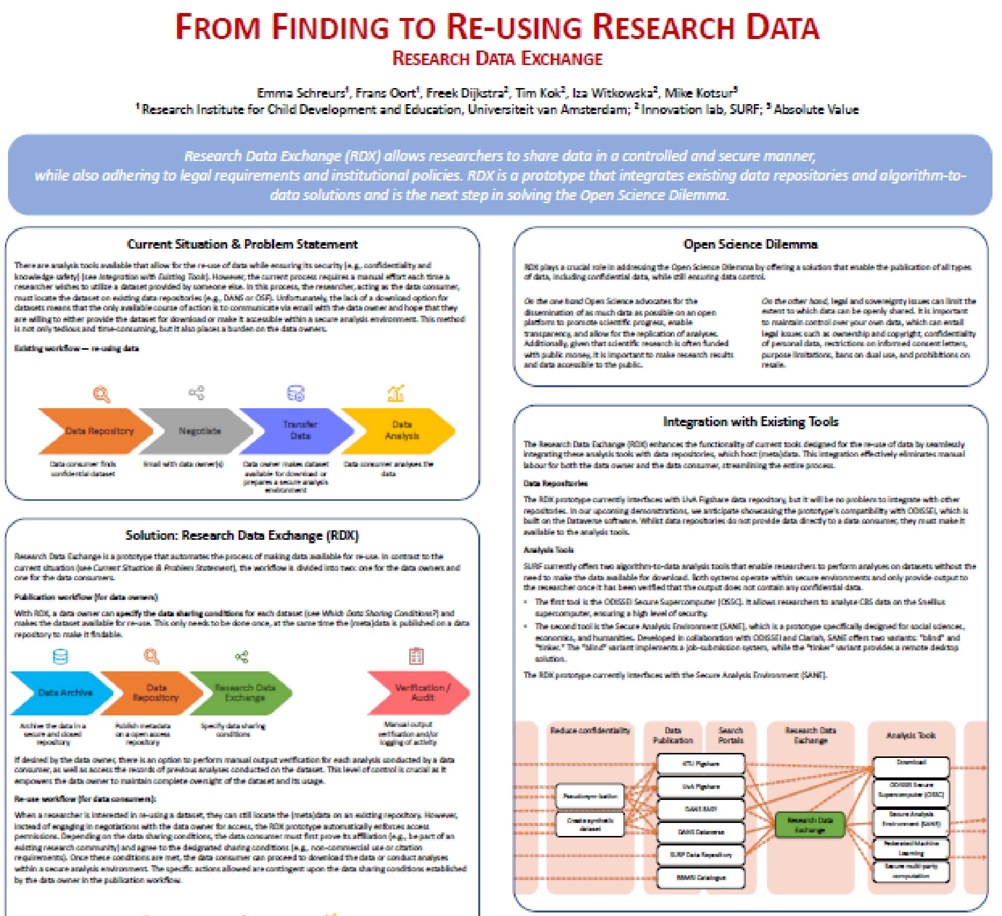
Towards a Modular Infrastructure for Comprehensive RDM[1]

Idea for a TDCC SSH Challenge Project[2]

Frans Oort, Eva Lekkerkerker, Emma Schreurs, Tako Horsley[3]

University of Amsterdam, May 2024

Dutch universities have policies or guidelines for RDM in the SSH domain.[4] However, not all universities have the complete infrastructure required for full implementation of RDM policies, including FAIR data policies. To ensure compliance with applicable laws, regulations and institutional policies, it is important that infrastructure is appropriate, compliant and easy to use, freeing researchers from administrative tasks, reducing the workload of support staff and providing accurate management information.



From Finding to Re-using Research Data

Research Data Exchange

Emma Schreurs[1], Frans Oort[1], Freek Dijkstra[2], Tim Kok[2], Iza Witkowska[2], Mike Kotsur[3]
[1] Research Institute for Child Development and Education, Universiteit van Amsterdam; [2] Innovation lab, SURF; [3] Absolute Value

Research Data Exchange (RDX) allows researchers to share data in a controlled and secure manner, while also adhering to legal requirements and institutional policies. RDX is a prototype that integrates existing data repositories and algorithm-to-data solutions and is the next step in solving the Open Science Dilemma.

Current Situation & Problem Statement

There are analysis tools available that allow for the re-use of data while ensuring its security (e.g., confidentiality and knowledge safety) (see *Integration with Existing Tools*). However, the current process requires a manual effort each time a researcher wishes to utilize a dataset provided by someone else. In this process, the researcher, acting as the data consumer, must locate the dataset on existing data repositories (e.g., DANS or OSF). Unfortunately, the lack of a download option for datasets means that the only available course of action is to communicate via email with the data owner and hope that they are willing to either provide the dataset for download or make it accessible within a secure analysis environment. This method is not only tedious and time-consuming, but it also places a burden on the data owners.

Existing workflow — re-using data

Data Repository
Data consumer finds confidential dataset
Negotiate
Email with data owner(s)
Transfer Data
Data owner makes dataset available for download or prepares a secure analysis environment
Data Analysis
Data consumer analyses the data

Solution: Research Data Exchange (RDX)

Research Data Exchange is a prototype that automates the process of making data available for re-use. In contrast to the current situation (see *Current Situation & Problem Statement*), the workflow is divided into two: one for the data owners and one for the data consumers.

Publication workflow (for data owners)

With RDX, a data owner can **specify the data sharing conditions** for each dataset (see *Which Data Sharing Conditions?*) and makes the dataset available for re-use. This only needs to be done once, at the same time the (meta)data is published on a data repository to make it findable.

Data Archive
Archive the data in a secure and closed repository
Data Repository
Publish metadata on a open access repository
Research Data Exchange
Specify data sharing conditions
Verification / Audit
Manual output verification and/or logging of activity

If desired by the data owner, there is an option to perform manual output verification for each analysis conducted by a data consumer, as well as access the records of previous analyses conducted on the dataset. This level of control is crucial as it empowers the data owner to maintain complete oversight of the dataset and its usage.

Re-use workflow (for data consumers):

When a researcher is interested in re-using a dataset, they can still locate the (meta)data on an existing repository. However, instead of engaging in negotiations with the data owner for access, the RDX prototype automatically enforces access permissions. Depending on the data sharing conditions, the data consumer must first prove its affiliation (e.g., be part of an existing research community) and agree to the designated sharing conditions (e.g., non-commercial use or citation requirements). Once these conditions are met, the data consumer can proceed to download the data or conduct analyses within a secure analysis environment. The specific actions allowed are contingent upon the data sharing conditions established by the data owner in the publication workflow.

Open Science Dilemma

RDX plays a crucial role in addressing the Open Science Dilemma by offering a solution that enable the publication of all types of data, including confidential data, while still ensuring data control.

On the one hand Open Science advocates for the dissemination of as much data as possible on an open platform to promote scientific progress, enable transparency, and allow for the replication of analyses. Additionally, given that scientific research is often funded with public money, it is important to make research results and data accessible to the public.

On the other hand, legal and sovereignty issues can limit the extent to which data can be openly shared. It is important to maintain control over your own data, which can entail legal issues such as ownership and copyright, confidentiality of personal data, restrictions on informed consent letters, purpose limitations, bans on dual use, and prohibitions on resale.

Integration with Existing Tools

The Research Data Exchange (RDX) enhances the functionality of current tools designed for the re-use of data by seamlessly integrating these analysis tools with data repositories, which host (meta)data. This integration effectively eliminates manual labour for both the data owner and the data consumer, streamlining the entire process.

Data Repositories

The RDX prototype currently interfaces with UvA Figshare data repository, but it will be no problem to integrate with other repositories. In our upcoming demonstrations, we anticipate showcasing the prototype's compatibility with ODISSEI, which is built on the Dataverse software. Whilst data repositories do not provide data directly to a data consumer, they must make it available to the analysis tools.

Analysis Tools

SURF currently offers two algorithm-to-data analysis tools that enable researchers to perform analyses on datasets without the need to make the data available for download. Both systems operate within secure environments and only provide output to the researcher once it has been verified that the output does not contain any confidential data.

- The first tool is the ODISSEI Secure Supercomputer (OSSC). It allows researchers to analyse CBS data on the Snellius supercomputer, ensuring a high level of security.
- The second tool is the Secure Analysis Environment (SANE), which is a prototype specifically designed for social sciences, economics, and humanities. Developed in collaboration with ODISSEI and Clariah, SANE offers two variants: "blind" and "tinker." The "blind" variant implements a job-submission system, while the "tinker" variant provides a remote desktop solution.

The RDX prototype currently interfaces with the Secure Analysis Environment (SANE).

Which Data Sharing Conditions?

a sharing conditions
liation
itions
alysis environment
lysis output before releasing the
orithms for analysis

Already implemented by the prototype

- Verify email address
- Sign data sharing conditions
- Download dataset
- Analyse in a secure analysis environment
- Verification of the analysis output before releasing the output

Objectives & Findings

s two primary objectives: to showcase the technology and to gain a deeper understanding of the roles , particularly the distinction between the researcher who generated the dataset and the data steward institute. One key aspect we aimed to explore was the verification of analysis outputs for data: formed by the researcher, who knows the ins-and-outs of the dataset, or by the data steward, who has mpliance.

both parties desire to retain control. However, for many datasets, it may not be necessary to verify ds if robust logging and monitoring mechanisms are in place and there is the option to audit past ementation of an automated output check could potentially aid in prioritizing these audits, this topic ation in future discussions.

Are You Interested?

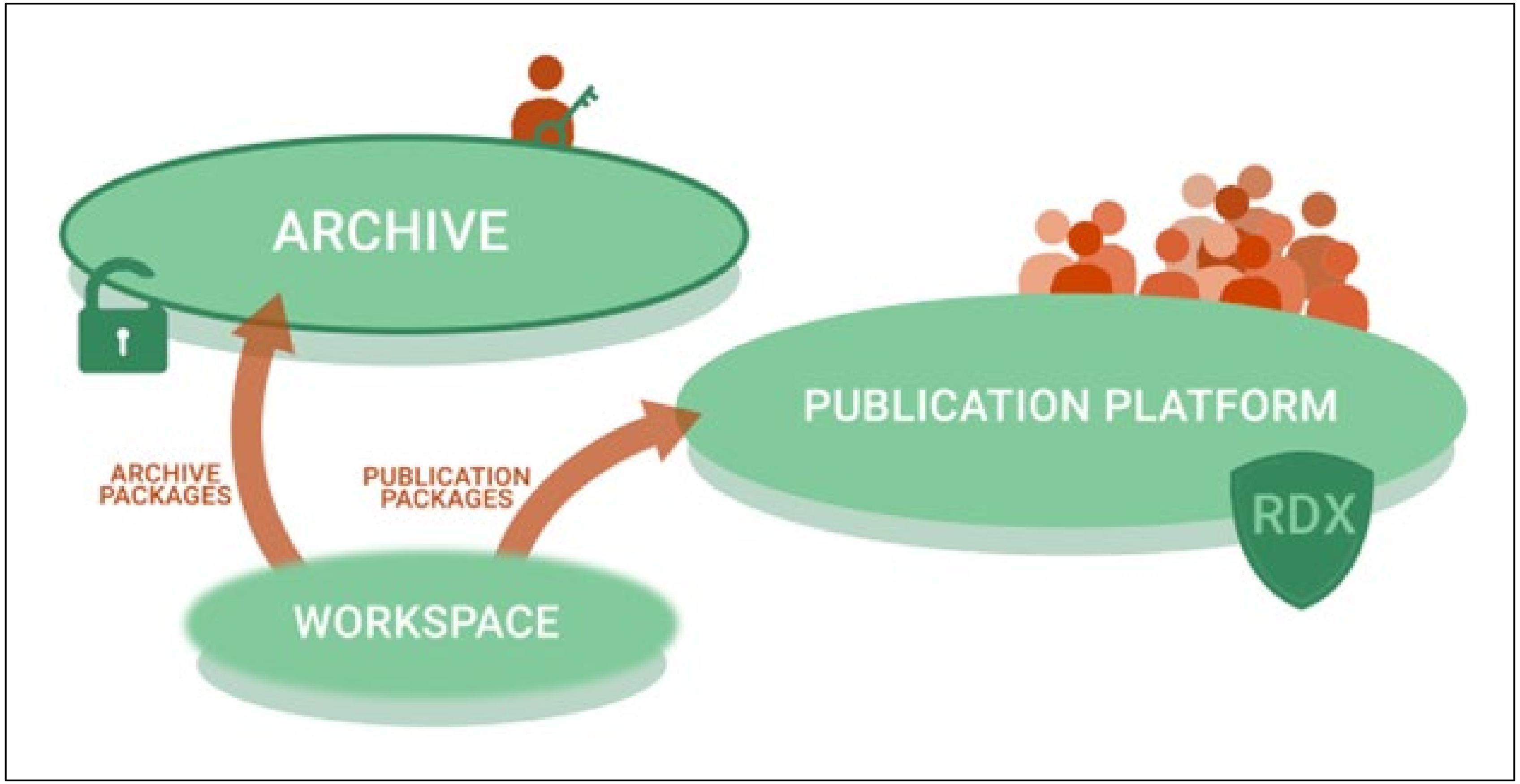
If you are a data owner or data steward with datasets that require specific use conditions, we would greatly appreciate your insights on the viability of this approach. Specifically, we would like to hear your perspective on the data sharing conditions that would enable you to make these datasets available for re-use, the feasibility of enforcing such conditions, and whether the logging and monitoring features provided are adequate for cases where enforcement may not be possible. Your input will help us further refine and enhance the system.

Please email freek.dijkstra@surf.nl to reach out to us!

The Research Data Exchange (RDX) is a use-case in the Amsterdam Data Exchange (AMdEX) project. The AMdEX project is sponsored the European Union and the Province of North-Holland.

SURF AMdEX European Union European Regional Development Fund Provincie Noord-Holland

A Modular Infrastructure for Comprehensive RDM



Storage, Archiving, Publication of Research Data

- Temporary workspace storage
- Closed storage for archiving (raw, sensitive) data
- Publication platform for publishing data
- A 'research data exchange' (RDX) for responsible data sharing
- A 'FAIR data hub' (FDH) for transferring data with essential metadata
- Data management software
- A 'research management services portal' (RMSP) with automated support and workflows

https://zenodo.org/records/11220690

Towards a Comprehensive OS Infrastructure

- EOSC as a federated 'system of systems'
- 'Systems' are storage, tools, services (and support, and communities)
- Focus on:
  - Rules of participation
  - Interoperability
  - Responsible information sharing (RDX)
  - Indexing and dissemination
- Inventory of existing 'systems', fit-gap analysis
- Design, build, develop, maintain the missing links and 'systems'

- Funding and governance:
  - Existing systems already have funding and the governance that goes with it
  - New systems need to be funded by the EC (i.e. the member states)
  - Governance by EOSC Association (members represent 'systems')
  - Mandated members advise national governments on funding, governance
  - Commercial parties participate as observers (to maintain digital sovereignty)

References

- On the Need to Establish Public Infrastructure to Preserve Digital Sovereignty: https://zenodo.org/records/8328514 (LERU Task Force, www.leru.org)
- Beyond APC: On the need for Diamond Open Access Publication Platforms: https://zenodo.org/records/4758335
- University Journals - Consolidating Institutional Repositories in a Digital Open Access Publication Platform: https://zenodo.org/records/3260292 (Liber Quarterly, 30, 2020, 1-15).
- Policy Principles for Research Data Management: https://zenodo.org/records/10954657
- Research Data Exchange: https://zenodo.org/records/8269273
- FAIR Data Hub: https://zenodo.org/records/11201003
- Towards a Modular Infrastructure for Comprehensive RDM: https://zenodo.org/records/11220690
- Data sovereignty, data governance and digital sovereignty: https://zenodo.org/records/10837008 (National Coordination Point Research Data Management, www.lcrdm.nl)