



Elly Dijk

Data Archiving and Networked Services - DANS

RDM support for researchers

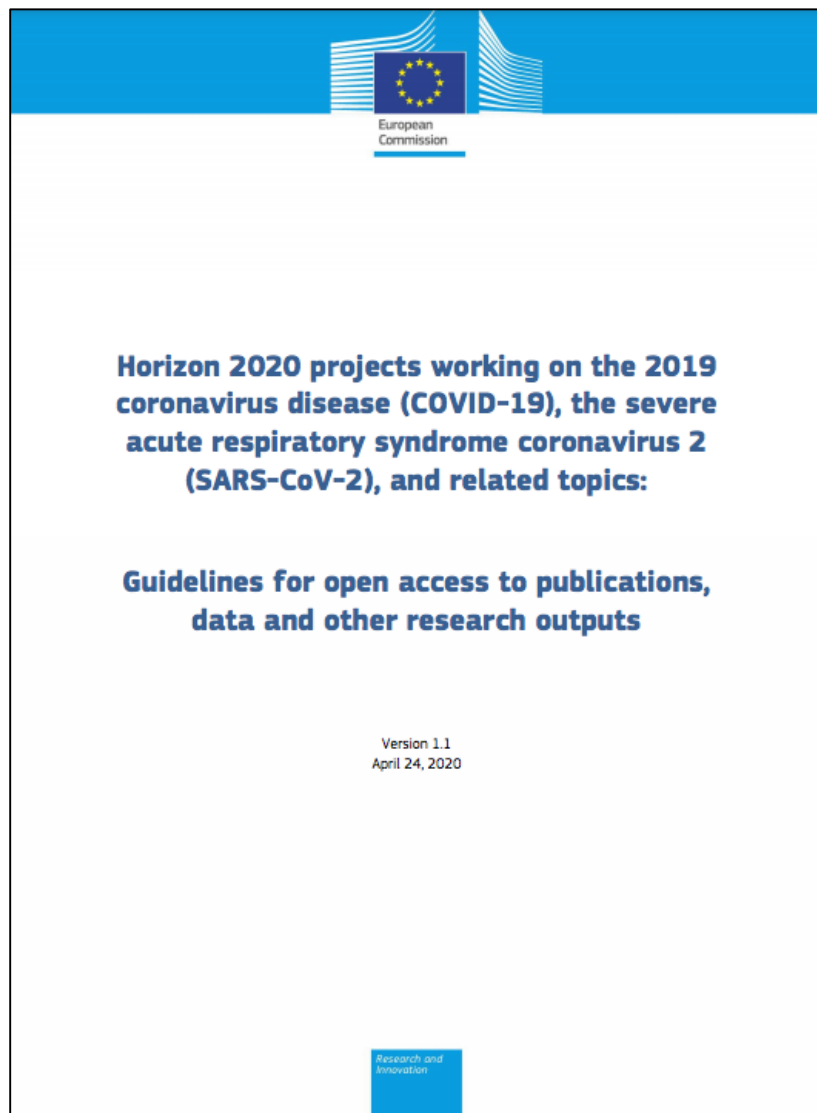
Output of the Task force Research Data Management



OpenAIRE-Advance GA | Online | 12 - 16 Oct 2020



European/International Policy promotes Open Science



Guidelines for open access ... April 2020



SE Guide RDM, 2018



SE framework discipline-specific RDM, 2018
Organisation for Economic Co-operation and Development (OECD), 2018



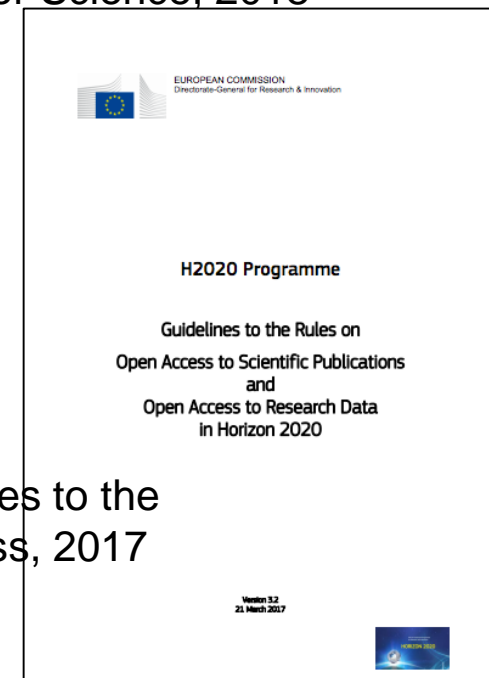
EOSC WG FAIR, 2018



EOSC Secretariat, March 2020



ICSU-International Council for Science, 2018

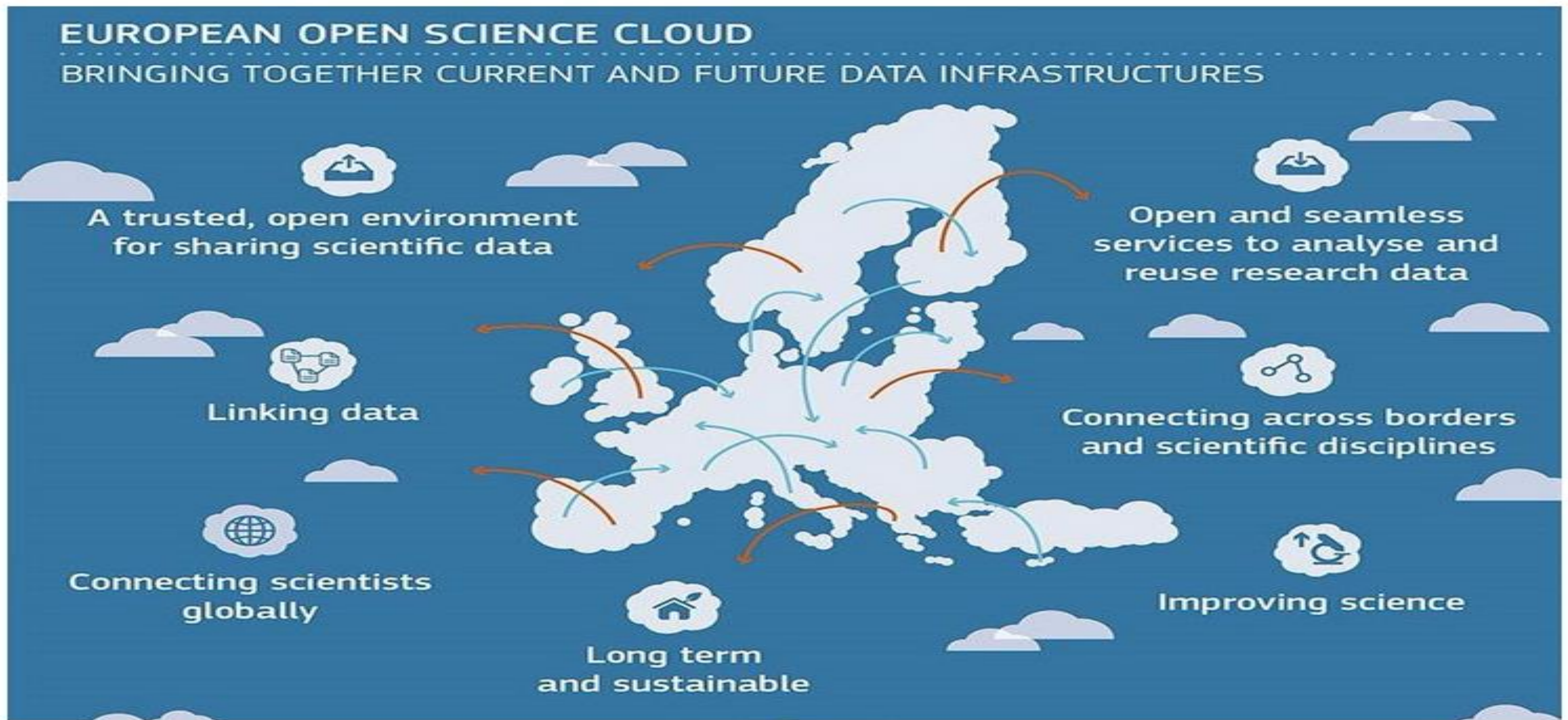


Horizon 2020 guidelines to the Rules on Open Access, 2017

OpenAIRE-Advance GA | Online | 12 - 16 Oct 2020



European Open Science Cloud - EOSC



EOSC Ecosystem – Turning FAIR into reality

FAIR =
Findable,
Accessible,
Interoperable,
Reusable



Task force RDM: Goal + activities

GOAL

Establish capacity and increase knowledge primarily among NOADs in order to support RDM activities, and support the Open Data Pilot/FAIR data/EOSC

HOW

Now: 31 members / 20 NOADs
Establishing working groups to examine different elements of the research data life cycle

HOW

Gathering examples of existing good practices

HOW

Finding gaps; developing new RDM materials

Output

Online guides, webinars, blog posts et cetera

Working groups

DMP resources – methodology for developing RDM roadmaps
Data reuse examples
The importance of long-term preservation
Use cases DMPs of existing projects
Survey: how do researchers in your institution manage their data?
Promotion of the task force output

OpenAIRE RDM Task force Outputs

ARGOS – DMP tool

Costs to manage and share data:
10.5281/zenodo.3837717

How to deal with non-digital data:
10.5281/zenodo.4057878

Webinar (with EOSC-hub): [Data Privacy and Sensitive Data Services](#), 6 Dec 2018

Blog: [Electronic Lab Notebooks, should you go 'e'?](#)

Check our National Open Access Desks Starter Kit:
<https://www.openaire.eu/rdm-noads-starter-kit>

Webinar on **Amnesia** - data anonymization tool:

<https://www.openaire.eu/item/amnesia>

Storing sensitive data: 10.5281/zenodo.4048403
Find a trustworthy repository: 10.5281/zenodo.4020812
Data formats for preservation: 10.5281/zenodo.4041512
Raw data, backup and versioning:
10.5281/zenodo.4041557



Expecting: a number of blog posts on data-reuse with data reuse examples

Identifiers to improve dissemination:
10.5281/zenodo.1051028
Managing access to sensitive data:

10.5281/zenodo.4048403

Webinar (with FREYA): [New](#)

[developments in the field of Persistent](#)

[Identifiers](#), 10 Jan 2019.

Ref: Research data lifecycle: UK Data

Archive

Electronic Lab Notebooks - should you go "e"?

scholarly communication infrastructure open science skills



Asger Væring Larsen



21 Dec 2018



5886 Hits



0 Comments

Autl

Lars

Which tool to choose?

A lab
note

exper

But,

note

enal

As n

Metl

for tl

A number of institutions have done comparison tests of different tools. In 2017, the Gurdon Institute tested 30 different tools and compared features – see their conclusions [here](#). They have listed a number of relevant questions to consider when choosing an ELN.

Also in 2017, Harvard conducted a comprehensive study into ELN's and have created a matrix comparing 29 tools on 50 different parameters. Their results can be seen [here](#).

In 2018, the University of Delft held a workshop on Digital Notebooks. Presentations from that workshop are available in [Zenodo](#) and a report on the workshop can be found [here](#). All of these investigations offer their experience openly for everyone to use, and share potential pros and cons of the available tools. It is apparent that the choice of tool depends on several different user needs like: ability to share protocols, compatibility with other tools, cloud storage, cost reduction, an Open Source solution, or capability of dictation/voice input.

So, what's all the hubbub about? Should a researcher switch from paper to electronic lab notebooks (ELN's)? Why? And where to get started? This blogpost intends to highlight the advantages of switching from paper notebooks to an electronic version of it, and to point researchers towards different initiatives and case studies that can facilitate the decision of adopting an ELN.

Guides

Data

What

THE CONTEXT

WHY IS IT
NECESSARY?

**HOW TO DEAL
WITH THIS?**

RESOURCES

How to deal with this?

As an example, in the biomedical imaging field, a realisation of the huge variety of file formats that exist led to an initiative to make these interoperable. As part of the OMERO project, Bioformats is a software plugin which allows the conversion of multiple established proprietary and standard file formats. Image analysis software such as ImageJ (free and open source) have adopted Bioformats as a plugin to allow users to read and write their image data without having to consider their origin. However, such tools may not always exist for different disciplines, and a researcher should consider storing their acquired data in a standard format at the earliest available opportunity. Many (most?) commercial and open source software packages allow conversion of data into standard formats and this should be exploited.

During the course of the digital revolution, a number of file formats have been recognised to be the file formats of choice for longevity and interoperability.

Please find below some useful links to resources about data formats for long-term storage:

- [Data description and formats](#). 4TU.Centre for Research Data
- [File formats](#). DANS - Data Archiving and Networked Services

As an example, the following table describes a variety of file formats for different disciplines that are either recommended or acceptable (from the UK Data Service):

Type of data

Recommended formats

Acceptable formats

negative consequence of making these data less interoperable.

Support

RESOURCES

[Open Science Primers](#)

[Guides](#)

[Factsheets](#)

[Use cases](#)

WHAT IS REQUIRED?

THE EC PREFERS CERTIFIED
REPOSITORIES

HOW TO COMPLY WITH THE
OPEN DATA PILOT
REQUIREMENTS?

WHICH REPOSITORY TO
USE?

SUPPORT ON METADATA,
SENSITIVE DATA, USAGE
LICENCES

ARE DATA PUBLISHING
COSTS SUPPORTED?

WHAT ARE THESE
REPOSITORY
CERTIFICATIONS BASED ON?

HOW CAN OPENAIRE HELP?

Which repository to use?

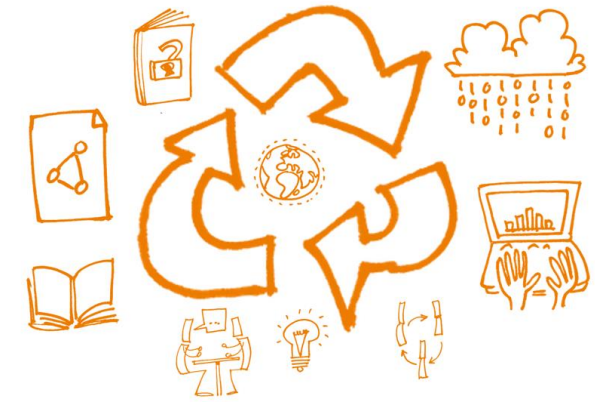
The general steps for finding a data repository are:

1. use a disciplinary repository if there is one;
2. alternatively, use the institutional repository, if you have one where the data will also be available for the long term;
3. use the catch-all repository [Zenodo](#), maintained by CERN;
4. or search the global [re3data.org portal](#) for a fitting repository - this provides several filtering options.

It's not easy to evaluate the quality of repositories, because this is influenced by many external factors, starting with the mission of the repository. For instance, does it explicitly aim for long-term preservation - with the appropriate expertise and budget - or not? Is it dedicated to a specific research community and familiar with their data formats, or is it generic? However, if you focus on repositories that are certified as being trustworthy, you simplify your selection process. So, if you don't have a disciplinary repository, and use the [re3data.org portal](#) for your search, we recommend that you filter on "Certificate" and look for the red icon (unfortunately, [OpenDOAR](#) has no such filter).



More RDM blog posts and other output



Blogs on Institutional RDM support:

<https://www.openaire.eu/blogs/use-cases-of-institutional-rdm-support-openaire-blog-series>

www.openaire.eu/blogs

Diagram: [Deposit your data in a data repository for long-term preservation](#): Why, When, How, What, Where

Expecting: DMP use cases project: overview of EC-DMPs (around 1,000) in repository of the [University of Vienna \(Phaidra\)](#)

What to deposit?

Is the data needed to reproduce your work?

Could this data be re-used?

Is the data unique?

Must it be kept as evidence or for legal reasons?

Should it be kept for its potential value?

Consider costs – do benefits outweigh cost?

Evaluate criteria to decide what to keep, because criteria may differ among projects

Reuse of data is only possible if they can be found in public places
([https://www.thelancet.com/pdfs/journals/lancet/PIIS1473-3099\(20\)30119-5.pdf](https://www.thelancet.com/pdfs/journals/lancet/PIIS1473-3099(20)30119-5.pdf))

Increased use and economic benefit

Criteria to decide what data to keep
(<https://www.data.cam.ac.uk/data-management-guide/looking-after-and-sharing-your-data#Preservation>)

Recorded factual material comm
scientific community as ne
(<https://www2.le.ac.uk/services/in-data>). Horizon2020 demand
documentation, tools and/or s

management

to do

with your

OpenAIRE guide and preferred file formats here.

Use domain-specific controlled vocabularies. Schema.org (<https://schema.org/>) is widely used to build controlled vocabularies, a more specific example is bioschemas.org (<https://bioschemas.org/specifications/DataSet/>); a collection of specifications that provide guidelines to facilitate a more consistent adoption of schema.org within the life sciences. Note: building a vocabulary is an advanced activity and no part of the regular research lifecycle.

Task force RDM - output



<https://www.openaire.eu/task-forces-in-openaire-advance>

Guides and other resources

- [How to find a trustworthy repository for your data?](#) (by Marjan Grootveld and Gültekin Gurdal)
- [Data formats for preservation](#) (by Paula Moura and S. Venkataraman)
- [How to deal with non-digital data?](#) (by Judit Fazekas-Paragh and S. Venkataraman)
- [How to deal with sensitive data?](#) (by Emilie Hermans and Olivia Kaiser)
- [RDM: Train the trainer resources](#) | Overview of Research Data Management (RDM) training materials for and by OpenAIRE NOADs
- [Raw data, backup and versioning](#) (by Paula Moura and S. Venkataraman)
- [How can identifiers improve the dissemination of your research outputs? Connect all your research products with your person identifier](#) (by Marjan Grootveld (DANS), Frances Madden (FREYA) and Alice Meadows (ORCID))
- [Indexing the RDM NOADs Starter Kit](#) (by Ádám Szalobagyi) *NB. log in needed to some resources*
- [Deposit your Data in a Data Repository for long-term preservation](#) (by Paula Moura, Iryna Kuchma and Alicia Gomez Sanchez). *NB. log in needed; press ctrl + or - for zoom in and out*
- [What will it cost to manage and share my data?](#) (Ryan O'Connor, Sarah Jones and Alexandra Delipalta)
- [FAQs - Questions from researchers in different countries about RDM](#) (check the FAQs tab!)

Webinars

- [RGPD y aspectos legales relacionados con la gestión de datos de investigación](#) - 23 June 2020
- [Amnesia - the OpenAIRE data anonymization tool](#) - 10 June 2020
- [GDPR and Research: Where do we stand?](#) - 4 May 2020
- [OpenAIRE Legal Policy Webinars](#) - 29 April 2020, 04 May 2020
- [Research Data Management and Legal issues related to research data](#) - 21 October 2019
- [Data stewardship e Research Data Management: Prerequisiti fondamentali per la scienza aperta](#) 10 May 2019
- [The role and value of data stewards in Universities: a TU Delft case study on data stewardship](#) 09 May 2019
- [FREYA and OpenAIRE: New developments in the field of Persistent Identifiers](#)
- [OpenAIRE - EOSC-hub webinar "Data Privacy and Sensitive Data Services"](#) - 06 December 2018
- [Research Data Management e politiche europee sui dati](#) - 15 November 2018

RDM task force blog posts

Blog posts

- [Electronic Lab Notebooks, should you go "e"?](#) (by Paula Martinez Lavanchy and Asger Væring Larsen)
- [4 Challenges for institutional research data management support](#) (by Emilie Hermans)
- [Research Data Management \(RDM\) Support at the University of Vienna](#) (by Susanne Blumesberger)
- [Turkey Research and Open Data Policy Task Force Established](#) (by Gültekin Gurdal)
- [Institutional RDM support at the University of Helsinki](#) (by Kimmo Koskinen)
- [Building bridges across Open Science in the UK: The RDA UK/OpenAIRE Advance Joint Workshop](#) (by Frank Manista)
- [Towards research data management support in Lithuania](#) (by Ieva Ceseviciute)
- [Institutional RDM Support at the Masaryk University \(Czech Republic\)](#) (by Michal Růžička, Jiří Marek, Marika Hrubá, Miroslav Bartošek)
- [Use cases of Institutional RDM Support - OpenAIRE blog series](#) (by Ellen Leenarts)

<https://www.openaire.eu/task-forces-in-openaire-advance>

Thank you!

Elly Dijk

elly.dijk@dans.knaw.nl

