



FAIR data and principles

Train-the-trainer Webinar for NI4OS project
February 19, 2020

 eosc-hub.eu

 [@EOSC_eu](https://twitter.com/EOSC_eu)





- Introduction
 - Moderators
 - Participants
 - Goal of the webinar
- Topics in relation to “FAIR data and principles”
 - FAIR data principles (René)
 - Services that support FAIR data (Cees)
 - Research data management (René)
 - NARCIS, the Dutch experience in creating a national catalogue of research output (Cees)
- Conclusion / discussion
 - Next steps
 - Feedback

René van Horik (rene.van.horik@dans.knaw.nl)

Data Archiving and Networked Services (DANS), The Netherlands
Trainer / project manager / researcher

Background: Humanities (PhD)



Cees Hof (cees.hof@dans.knaw.nl)

Data Archiving and Networked Services (DANS), The Netherlands
Trainer / project acquisition manager / DANS liaison life sciences

Background: Biology / Life Sciences (PhD)



Both active in RDM Training Work package of EOSC-hub project

- **Essentials for Data Support**

- <https://datasupport.researchdata.nl/en/>
- Online
- Face-to-Face 3 x per year

Essentials 4 Data Support is an introductory course for those people who (want to) support researchers in storing, managing, archiving and sharing their research data.

Essentials 4 Data Support is a product of Research Data Netherlands.

- **Delivering Research Data Management Services**

- <https://www.futurelearn.com/courses/delivering-research-data-management-services>
- Online
- Starts February 24th

What topics will you cover?

- What research support services are commonly offered and who is responsible for providing them?
- How to perform your own gap analysis and what tools you will need to make it happen
- How to support researchers to develop Data Management Plans (DMPs) and offer feedback or consultations
- How to provide for short and long term preservation and storage of research data and where possible, use existing community standards
- How to develop your own RDM Roadmap

EOSC-hub mobilises providers from 20 major digital infrastructures, EGI*, EUDAT CDI** and INDIGO-DataCloud jointly offering services, software and data for advanced data-driven research and innovation.

- * EGI is not an acronym (any more)
- ** CDI – Collaborative Data Infrastructure

E.g. Training on “research data management”

E.g. Webinar “T-t-T on FAIR data and principles” (NI4OS)

DANS is about *keeping* data FAIR





- To **inform** trainers on state-of-art of topics in relation to research data management
- To **provide** participants with training material on the topic (online source / practical experience /)

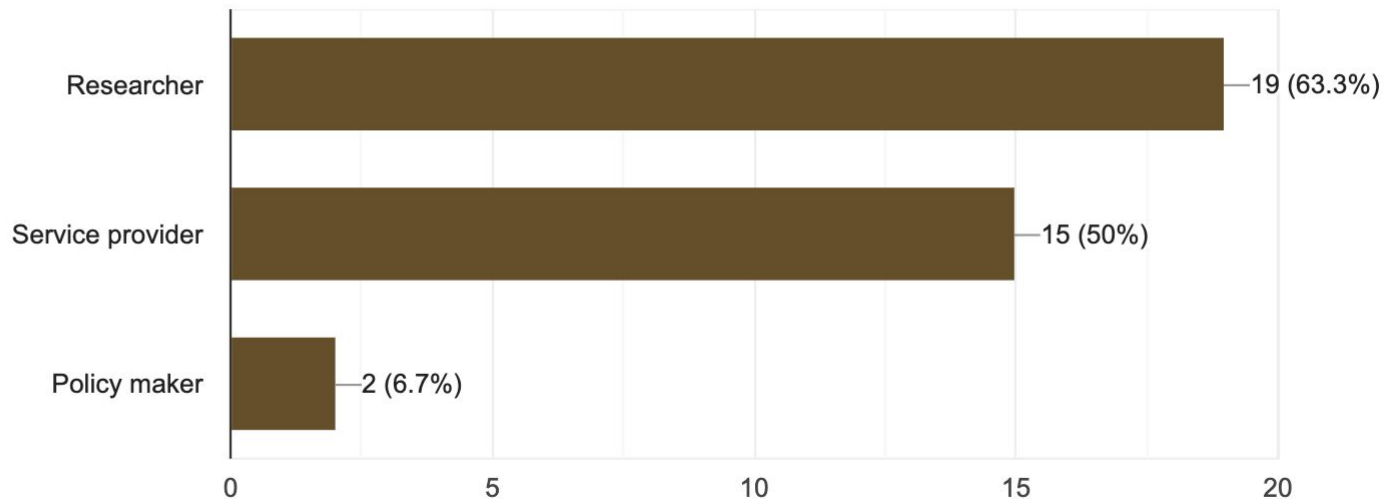
In order:

To assist participants in the webinar to organise trainings



Which of the following role suits you most?

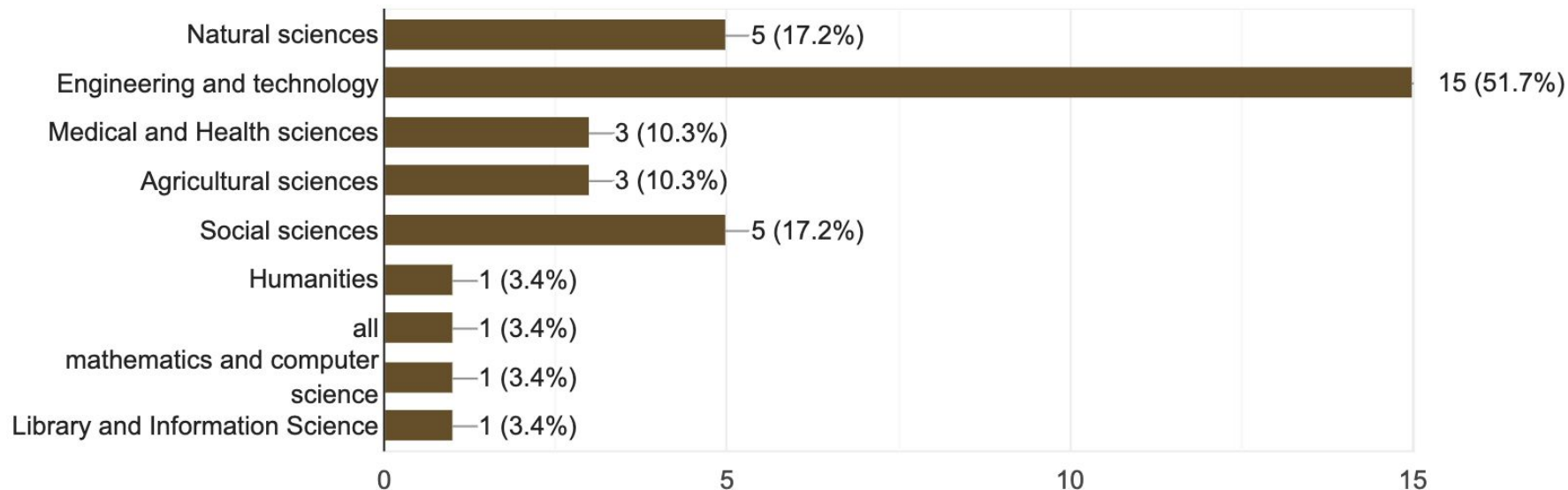
30 responses



In which scientific discipline are you active?



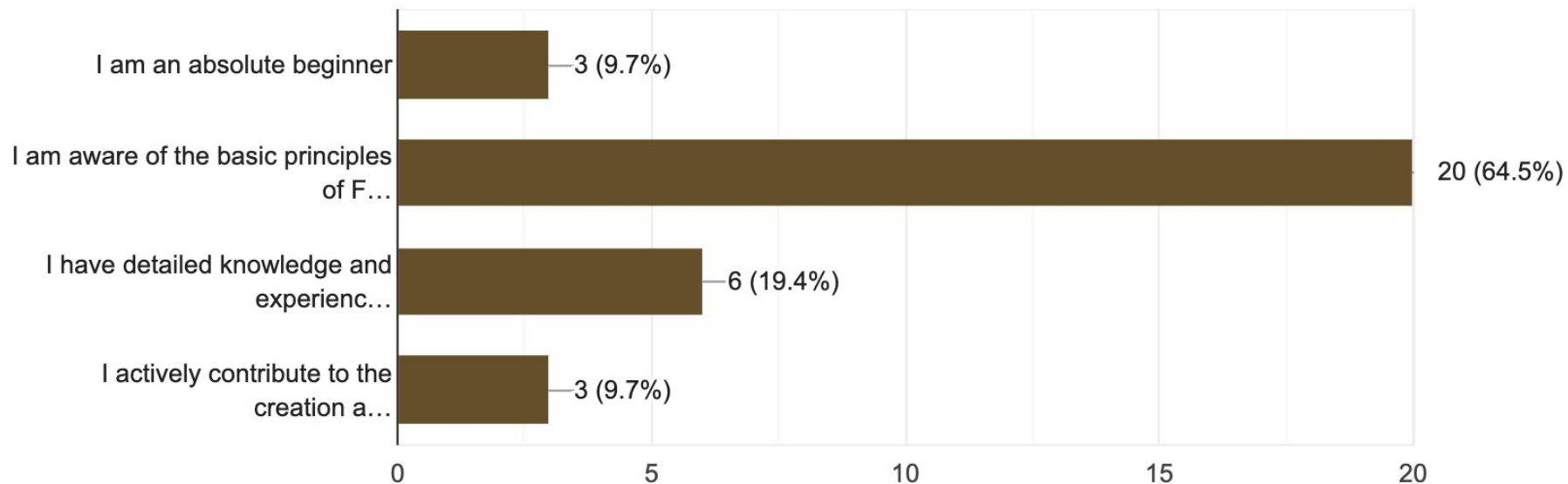
29 responses



How familiar do you consider yourself with FAIR data and / or open science?

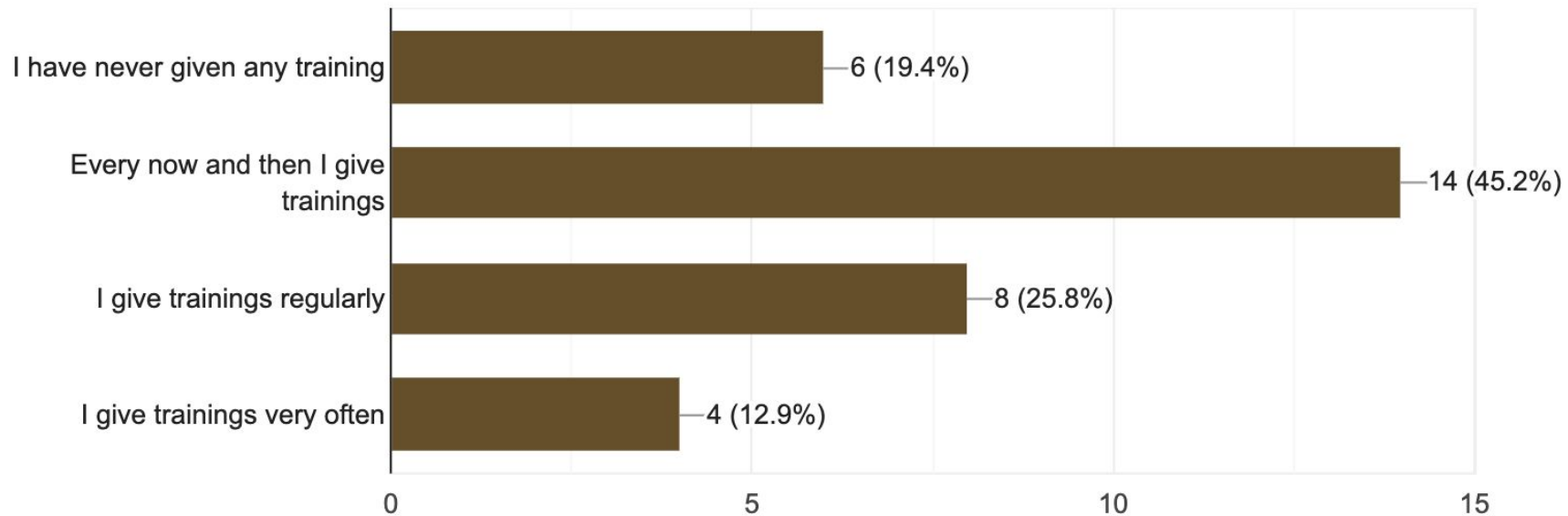


31 responses

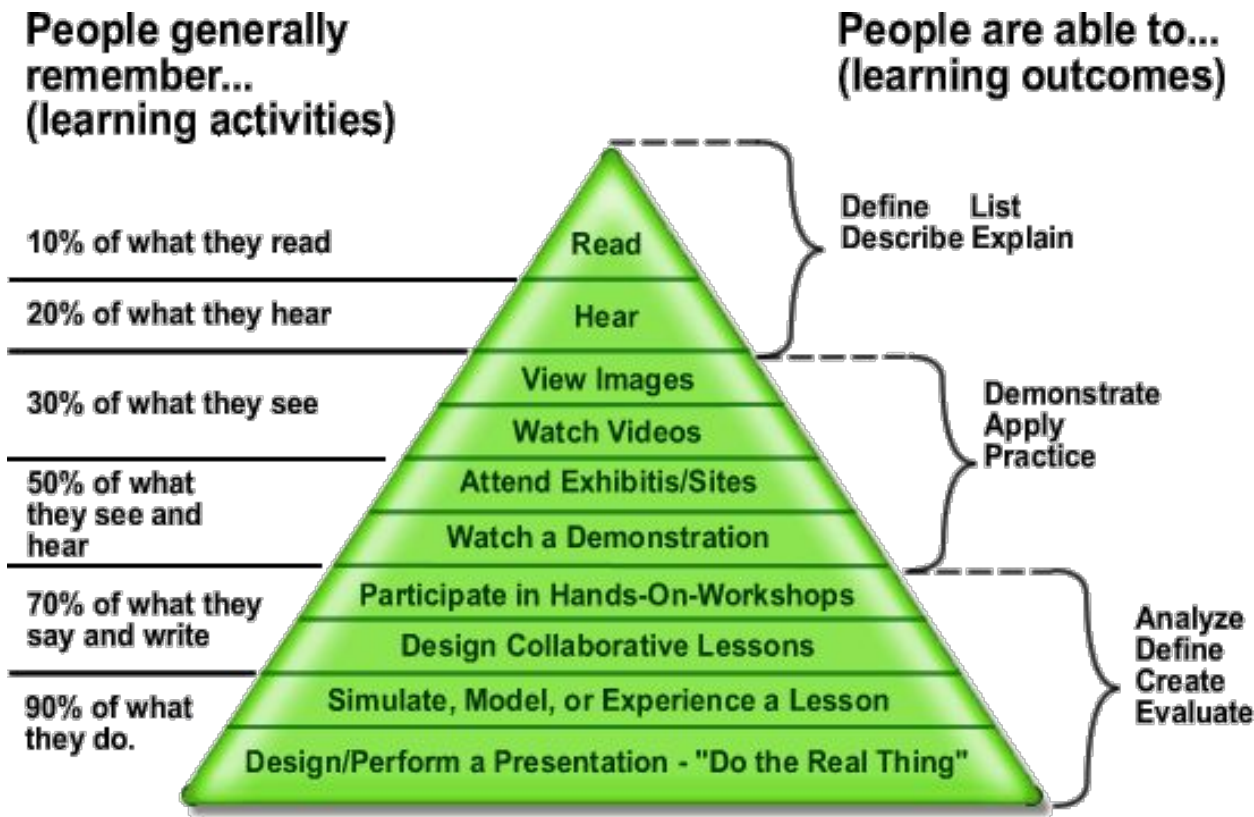


What is your experience with offering training?

31 responses



- Target group
- Learning goal
- Training format / material





European Research Infrastructures are important in facilitating open science.
Which RIs do you know? And in which scientific discipline?

Use the “shared notes” facility of the webinar training room...

1. FAIR data principles (René)

- Introduction
- Available training resources / experiences

2. Services that support FAIR data (Cees)

- Introduction
- Available training resources / experiences

3. Research data management (René)

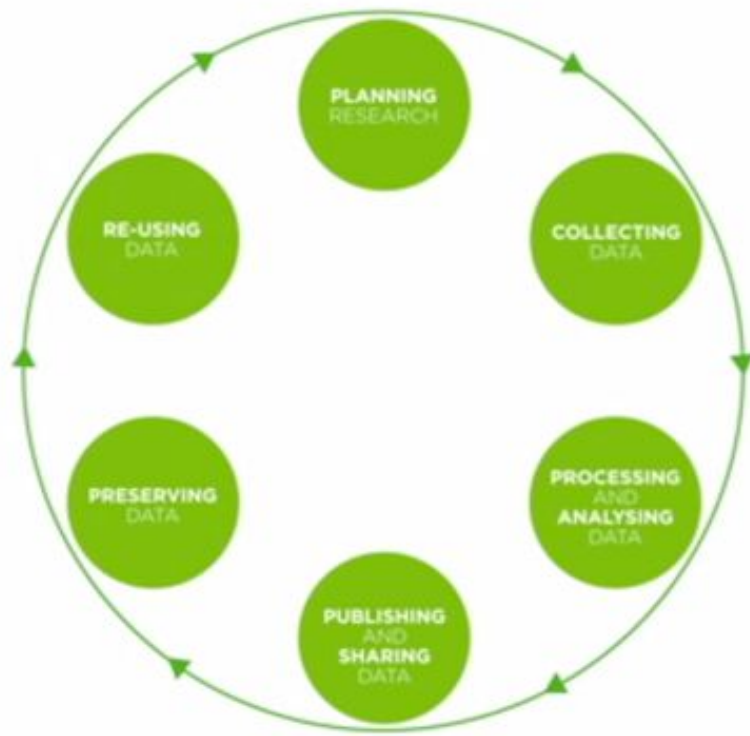
- Introduction
- Available training resources / experiences

4. NARCIS, the Dutch experience in creating a national catalogue of research output (Cees)

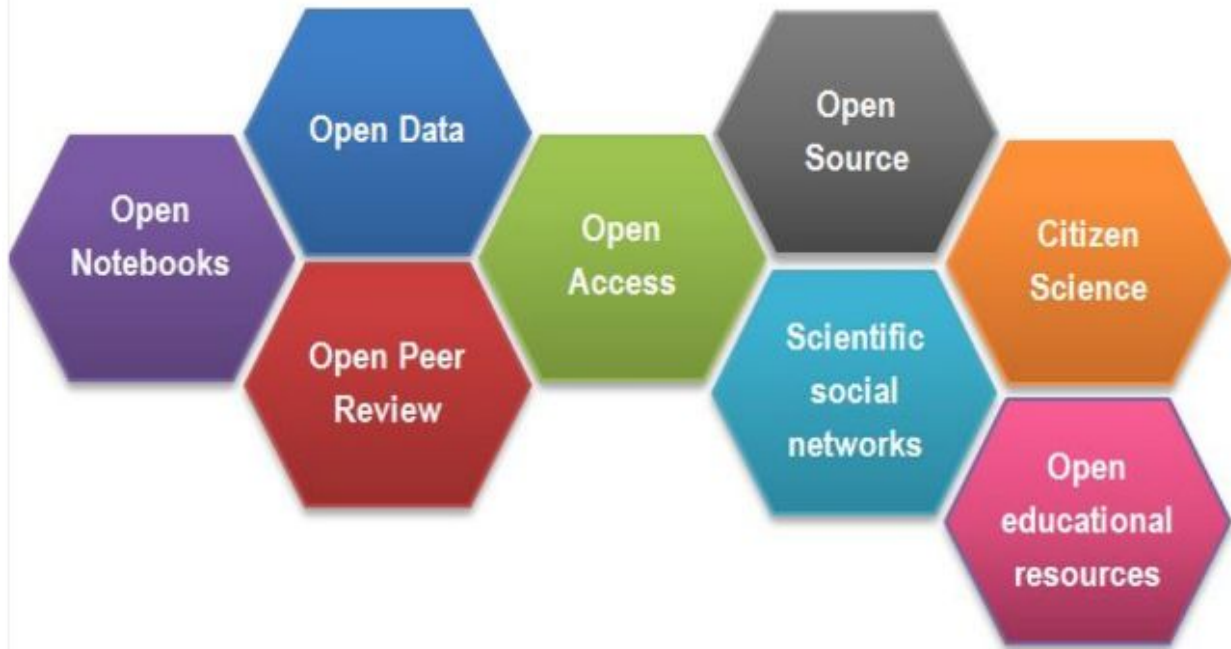
- Introduction
- Available training resources / experiences

Two basic concepts of the current research data landscape, that should be part of any training on FAIR data and principles:

- Research Data Lifecycle
- Open Science



Open Science is the movement to make scientific research (publications, software, data) and its dissemination accessible to all levels of society. Open Science is transparent and accessible knowledge that is shared and developed through collaborative networks.

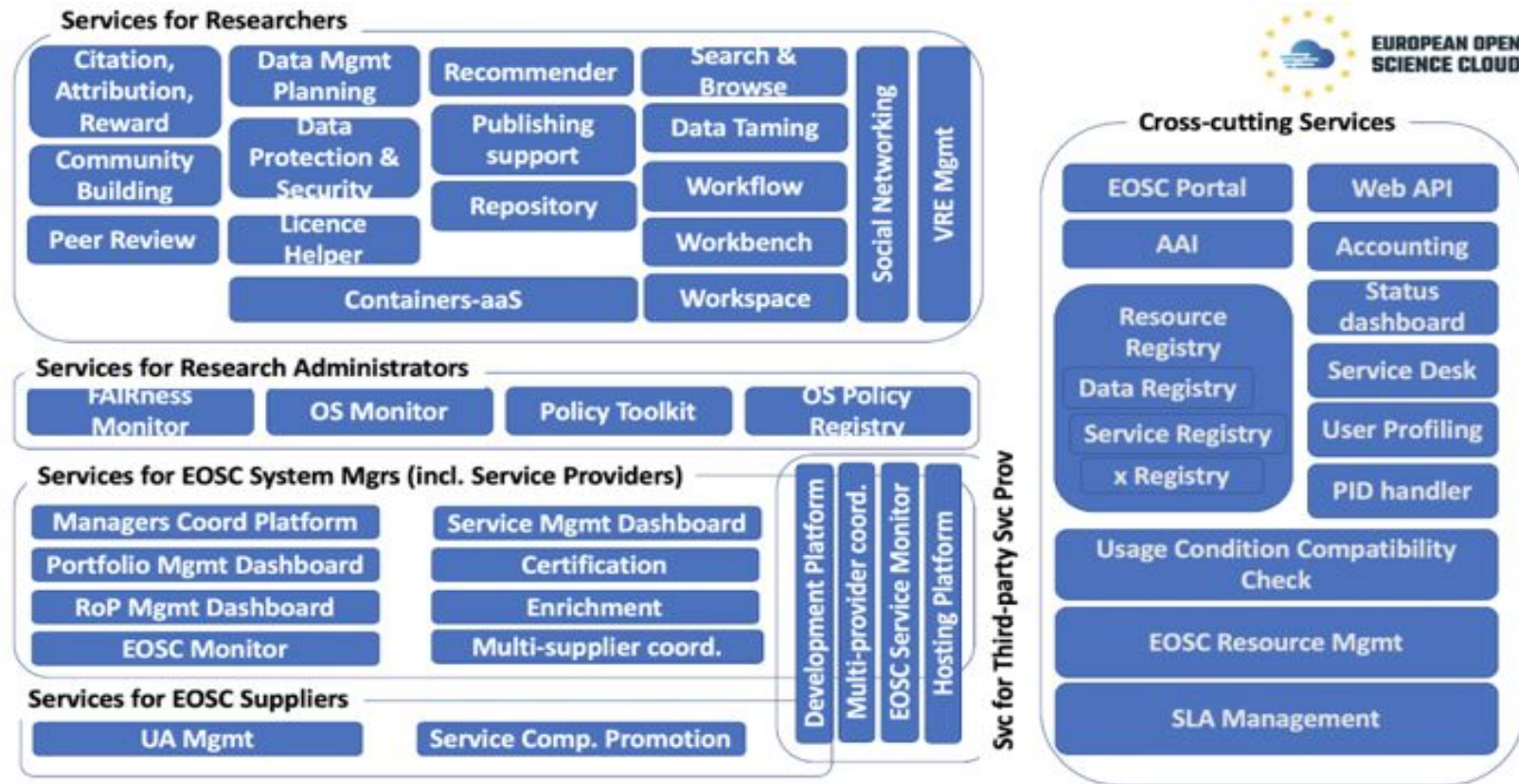




EUROPEAN OPEN SCIENCE CLOUD

The **European Open Science Cloud (EOSC)** is a [European Commission](#) initiative aiming at developing an infrastructure providing its users with services promoting [open science](#) practices. Besides being open science oriented, the envisaged infrastructure is built by aggregating services provided by several providers following a [System of systems](#) approach.

The initiative started in 2015 with the plan that its organizers finish it by 2020.^[1] A European Union committee on research endorsed a plan for the cloud's development in May 2018.^[2] The European Open Science Cloud officially launched in November 2018, starting to provide access to services via their [EOSC Portal](#)^[3].





The future of science is Open

START YOUR RESEARCH
TRAINING NOW

USE FOSTER TO:



Access Courses

Put Open Science into practice with our [Open Science training toolkit](#). Our [courses](#) are authored by experts and experienced educators.



Earn Badges

Get recognised for taking [our courses](#) and follow our [learning paths](#) to specialisations.



Participate

Join our community of trainers and access our [Trainers' Corner](#).



Promote Open Science

Use the [Open Science training handbook](#). In a variety of formats and languages.

FAIR data principles

René van Horik (DANS)



EOSC-hub

FAIR data principles

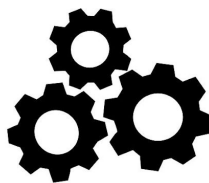
F
indable



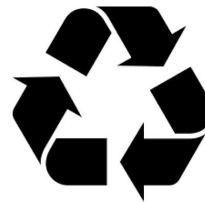
A
ccessible



I
nteroperable



R
eusable



Source: https://en.wikipedia.org/wiki/FAIR_data#/media/File:FAIR_data_principles.jpg

1. **Findable** – Easy to find by both humans and computer systems and based on mandatory description of the metadata that allow the discovery of interesting datasets;
2. **Accessible** – Stored for long term such that they can be easily accessed and/or downloaded with well-defined license and access conditions (Open Access *when possible*), whether at the level of metadata, or at the level of the actual data content;
3. **Interoperable** – Ready to be combined with other datasets by humans as well as computer systems;
4. **Re-usable** – Ready to be used for future research and to be processed further using computational methods.

source : <http://www.nature.com/articles/sdata201618>
www.force11.org/group/fairgroup/fairprinciples





Detailed description of FAIR principles

<https://www.go-fair.org/fair-principles/>

Example of site where you can find more information on FAIR

<https://www.go-fair.org/resources/>

- [GO FAIR Materials](#)
- [GO FAIR Workshop Series](#)
- [FAQ](#)
- [Starter Kit for Research Data Management](#)
- [More on FAIR](#)
- [Glossary](#)

FAIRsFAIR (Fostering FAIR data practices) - www.fairsfair.eu

(aims to supply practical solutions for the use of the FAIR data principles throughout the research data life cycle)

GO FAIR initiative - www.go-fair.org

(aims to implement the FAIR data principles)

FAIRsharing - fairsharing.org


(A curated , informative and educational resource on data and metadata standards inter-related to databases and data policies)

Example: Tool to assess the FAIRness of a dataset


SATIFYD

Self-Assessment Tool to Improve the FAIRness of Your Dataset

Welcome to SATIFYD: the DANS Self-Assessment Tool to Improve the FAIRness of Your Dataset. This tool will show you how FAIR (Findable, Accessible, Interoperable, Reusable) your dataset is and will provide you with tips to score (even) higher on FAIRness. Ideally, you use this tool prior to the deposit in EASY.

The 12 questions touch upon the FAIR data principles  but do not strictly follow them. While answering the questions, the score per letter will be displayed underneath each letter. The more 'blue' the letters get, the more FAIR your dataset is. An overall score is provided at the end of the page.

Some questions are posed more than once (e.g. on metadata and data standards or usage licences), because the topics are relevant in more than one letter.

Want to know more? Please click *here* 

If you have any questions, please let us know by sending an e-mail 

Source: <https://satifyd.dans.knaw.nl/>

Services that support FAIR data

By Cees Hof (DANS)



Introduction

- Short general introduction
- Services that support FAIR data before & after research
- Services that support FAIR data during research

Tips on training and training resources

- RDM & FAIR options
- FAIR assessment tools

Practical training experiences

- Experiences from the field...



- We will touch upon a broad spectrum of services that can help researchers and datastewards creating FAIR data
- Examples will be provided across disciplines
- Focus on generic, low threshold services (not too technical, not commercial)
- For trainers that start to work in this field...

For train the trainers we are pionering!



<https://dictionary.casrai.org/>

Services – A function that is being executed on request that delivers certain expected results.

So basically everything that works for you.....



<https://dictionary.casrai.org/>

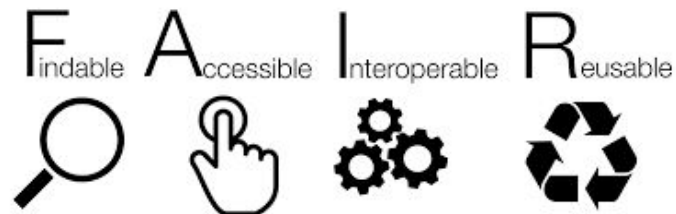


Use the CASRAI site for the consequent use of your terminology

Services – A function that is being executed on request that delivers certain expected results.

So basically everything that works for you.....

Services that support FAIR data means everything that helps you to increase the **F**indability, **A**ccessibility, **I**nteroperability or **R**eusability of your data.... and that is a lot!



“Recommendations for Services in a FAIR data ecosystem.”

Report by: FAIRsFAIR, RDA Europe,
OpenAIRE, EOSC-hub, FREYA

<https://zenodo.org/record/3585742#.XkWfohNKj0Q>



**Go read this
report!**



Recommendations for Services in a FAIR data ecosystem.

Report by: FAIRsFAIR, RDA Europe, OpenAIRE, EOSC-hub, FREYA

➡ The **biggest gap** at the moment: **Lack of a sustainable ecosystem** of independent interoperable services with governance, business model(s) and shared responsibilities to support the creation of FAIR research outputs.

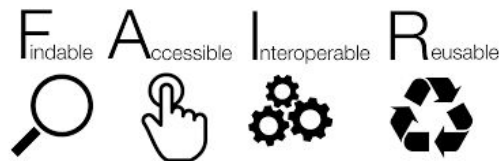
Recommendations for Services in a FAIR data ecosystem.

Report by: FAIRsFAIR, RDA Europe, OpenAIRE, EOSC-hub, FREYA

➔ The **biggest gap** at the moment: **Lack of a sustainable ecosystem** of independent interoperable services with governance, business model(s) and shared responsibilities to support the creation of FAIR research outputs.

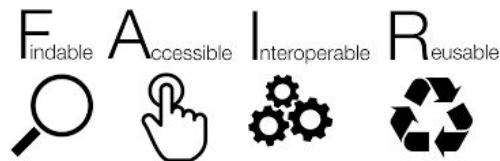


**Be prepared,
this means a
lot of work for
trainers...**



Distinguish 1:

- Need to address: the principles related to **findability** and **accessibility** which requires mostly **technical expertise** that can be addressed by generic services (e.g. **PIDs**, **cataloguing**, **discovery** and **storage**);
- Need to address: the principles related to **interoperability** and **reuse** which require services that cater to **disciplinary needs** with specific domain expertise (e.g. **ontologies**, **curation** and **stewardship** provided by domain repositories).



Distinguish 2:

- Services that can help you **prior to your research**, like Research Data Management services & Data Management Planning services.
- Services that help you **during research**, like Knowledge Organisation Systems (KOS) or PID services.
- Services that help you **after your research** like FAIR assessment tools, or data repositories with specific services.



- **RDM & DMP training**

Train your trainers about FAIR data and FAIR data services through Research Data Management (RDM) training and Data Management Planning (DMP's).

Most **online** and **face to face** training modules contain elements on FAIR services!

Essentials 4 Data Support:

<https://datasupport.researchdata.nl/en/>

Free online version, also face 2 face possibilities.



The MOOC: Delivering Research Data Management Services

See the teaser:

<https://www.futurelearn.com/courses/delivering-research-data-management-services>

A new Mooc starts on **February the 24th**.



Coursera: Data Management

<https://www.coursera.org/learn/data-management>

Not for free, several enrollment options.

For more info: Digital Curation Training pages of the DCC or the OpenAIRE Training pages

<http://www.dcc.ac.uk/training>

<https://www.openaire.eu/frontpage/webinars>

(Prior to research)

Coursera: Data Management

<https://www.coursera.org/learn/data-management>

Not for free, several enrollment options.

Trainers
TIP!

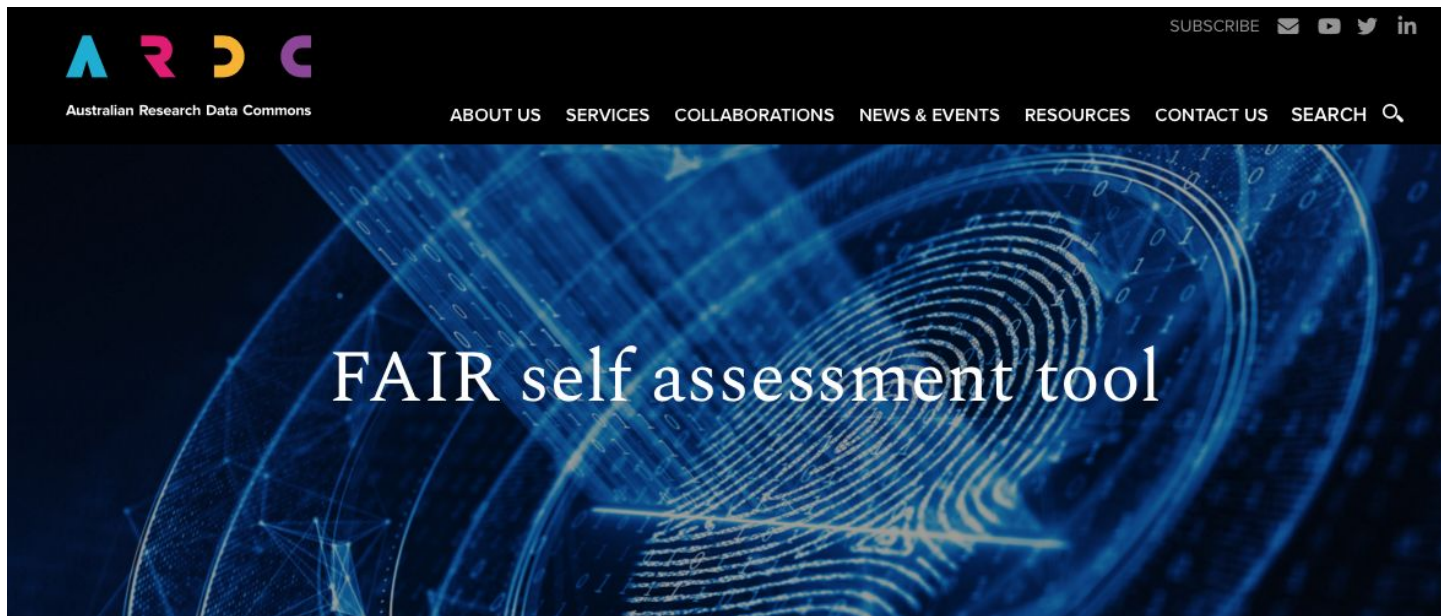
Obviously.. Try one or several of these courses before spreading the information to your trainers!

For more info: Digital Curation Training pages of the DCC or the OpenAIRE Training pages

<http://www.dcc.ac.uk/training>

<https://www.openaire.eu/frontpage/webinars>

FAIRdata assessment services:



<https://ardc.edu.au/resources/working-with-data/fair-data/fair-self-assessment-tool/>

- **Training on FAIR data assessment**

Checklist “How FAIR are your data?”

“A Checklist produced for use at the EUDAT summer school to discuss how FAIR the participant's research data were...”

<https://zenodo.org/record/1065991#.Xkz6rS2ZOL4>

FAIRdat tool

“Using this tool you will be able to score the 'FAIRness' of a dataset.”

<https://www.surveymonkey.com/r/fairdat>

FAIR enough? Checklist to evaluate FAIRness of data(sets)

Provided by DANS

<https://docs.google.com/forms/d/e/1FAIpQLSf7t1Z9IOBoj5GgWqik8KnhtH3B819Ch6lD5KuAz7yn0lOOpw/viewform>

- **Training on FAIR data assessment**

Checklist “How FAIR are your data?”

“A Checklist produced for use at the EUDAT summer school to discuss how FAIR the participant's research data were...”

<https://zenodo.org/record/1065991#.Xkz6rS2ZOL4>

FAIRdat tool

“Using this tool you will be able to score the 'FAIRness' of a dataset.”

<https://www.surveymonkey.com/r/fairdat>

FAIR enough? Checklist to evaluate FAIRness of data(sets)

Provided by DANS

<https://docs.google.com/forms/d/e/1FAIpQLSf7t1Z9IOBoj5GgWqik8KnhtH3B819Ch6lD5KuAz7yn0l0Opw/viewform>



These are all typical training tools/services (not production)

- **Training on FAIR services**

Major categories are:

- Choosing the proper **Persistent Identifiers** (PID)
- The use of **Knowledge Organisation Systems** (KOS)
- Choosing the right **data platform / data repository**

- **Training on FAIR services**



Training on these subjects will be very domain specific, choose the right trainers!

Major categories are:

- Choosing the proper **Persistent Identifiers** (PID)
- The use of **Knowledge Organisation Systems** (KOS)
- Choosing the right **data platform / data repository**

- **Training on Persistent Identifiers**

Very much under construction!



Training on PIDs can be very domain specific, choose the right trainers!

As a train the trainer:

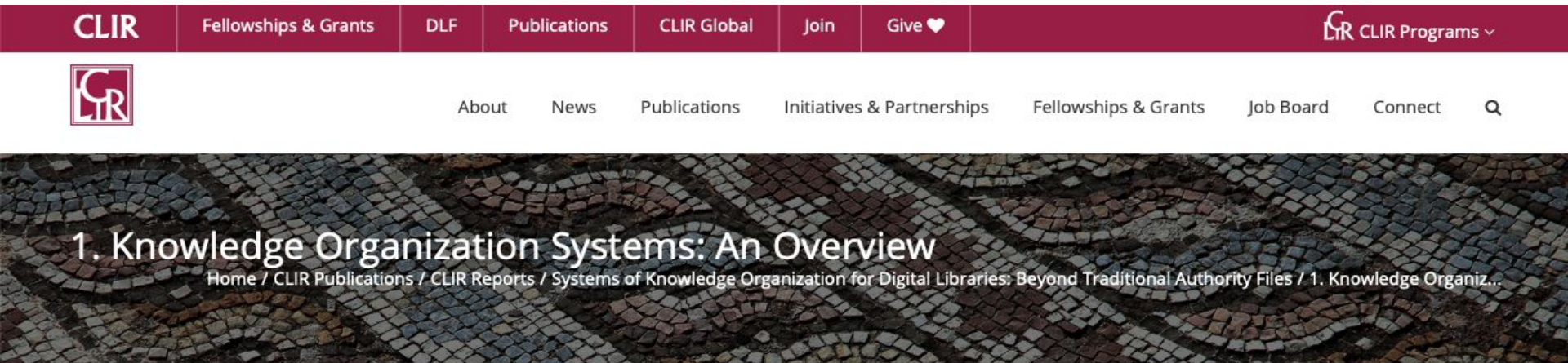
Get involved in the FREYA PID Forum!

<https://www.pidforum.org>

The PID Forum



- **Training on Knowledge Organisation Systems (KOS)**

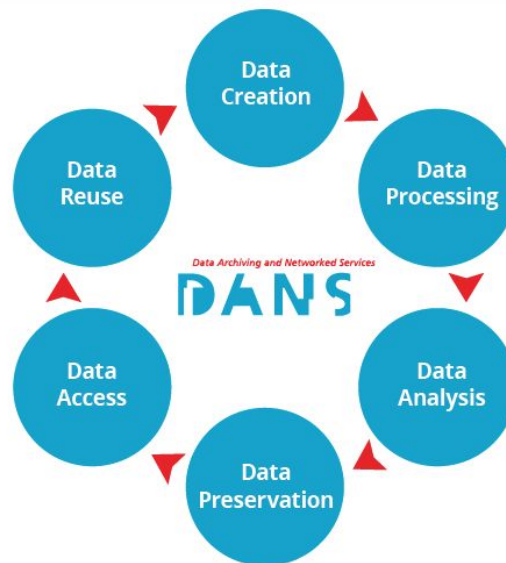


As an introduction start with the website of the Council on Library and Information Resources (CLIR):
<https://www.clir.org/pubs/reports/pub91/1knowledge/>

- **Training on Knowledge Organisation Systems (KOS)?**

Be inventive!

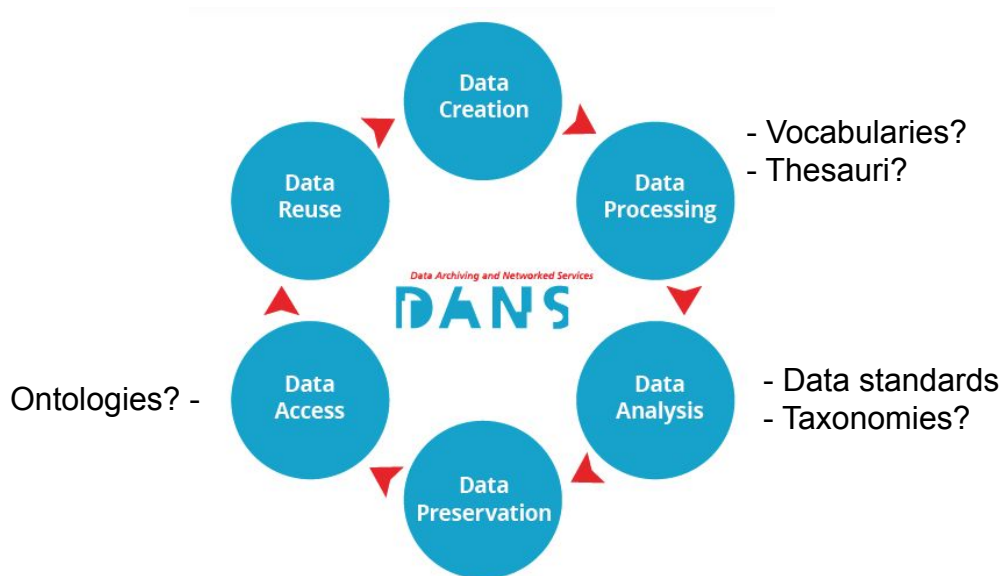
For example, develop an exercise around KOS services for stages of ongoing research....



- **Training on Knowledge Organisation Systems (KOS)?**

Be inventive!

For example, develop an exercise around KOS services for stages of ongoing research....



- **Training on data platforms and repositories**

Choosing the right platform / repository very much determines the FAIRness of the stored data!

Exercises / training around re3data.org are often clarifying the issues that require attention (metadata, certification, etc.) but **not readily available...**

<https://www.re3data.org>



**Thank you
for your attention!**

Questions?



EOSC-hub

 eosc-hub.eu  [@EOSC_eu](https://twitter.com/EOSC_eu)



This material by Parties of the EOsc-hub Consortium is licensed under a Creative Commons Attribution 4.0 International License.

Research Data Management

René van Horik (DANS)

Research data management refers to the development, execution and supervision of (research) plans, policies, programs and practices that control, protect, deliver and enhance the value of (research) data and information assets.

WHY IS IT IMPORTANT:

- Saves time and resources
- It helps to prevent errors and increases quality of research
- Allows to validate and replicate findings
- Facilitates sharing of data
-

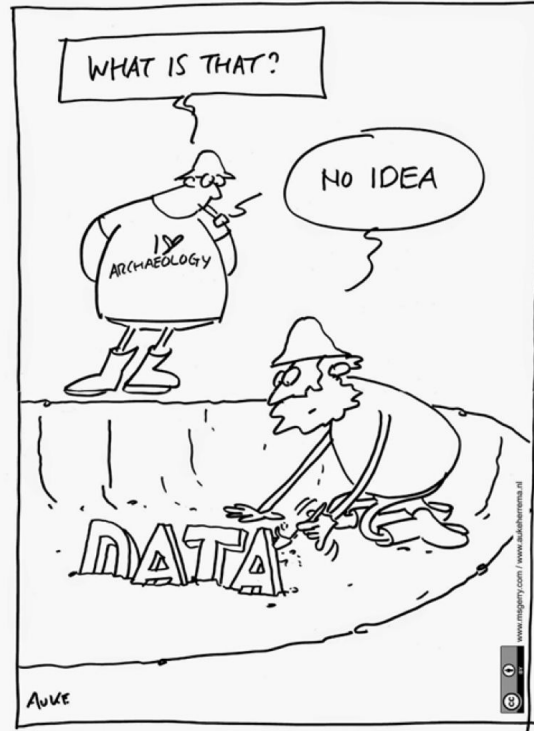




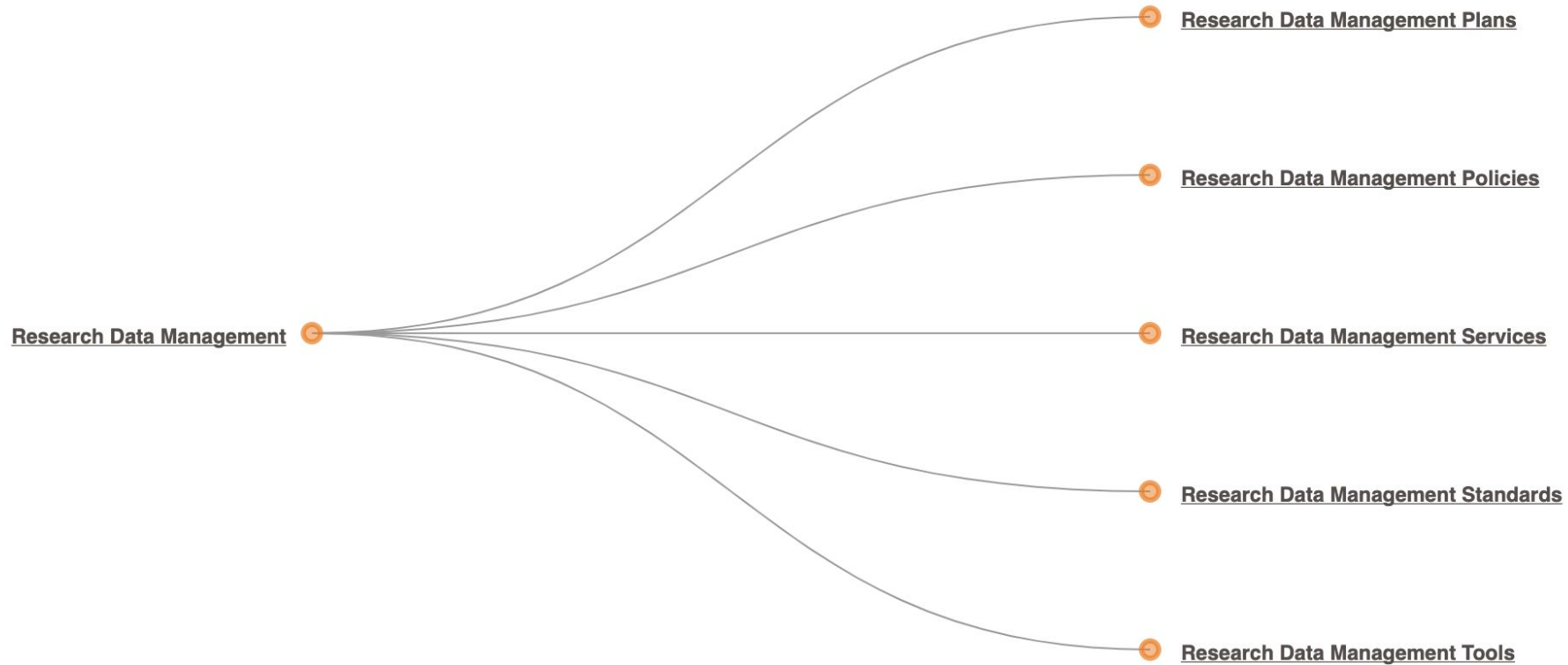
PUBLICATIONS AND DATA



THE FUTURE?



DATA FOR FUTURE GENERATIONS

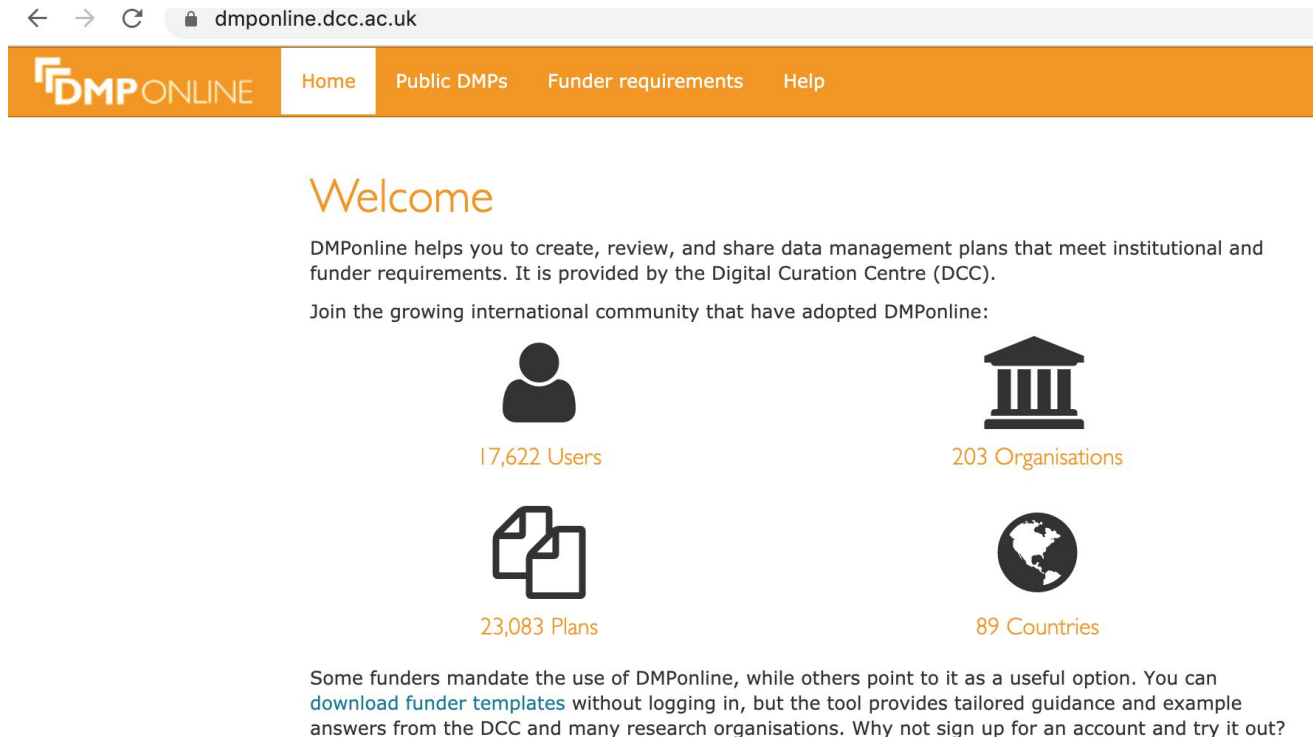


source: <https://www.fosteropenscience.eu/resources#tab-9>

- DMP = Document that contains information about handling, organising, documenting and enhancing research data, and enabling their sustainability and sharing for a research project
- A DMP Describes and analyzes workflows along the Research Data Lifecycle
- A DMP can be a few paragraphs short up to several pages long

ANNEX 1: DMP CORE REQUIREMENTS

1. **Data description and collection or re-use of existing data**
 - a. How will new data be collected or produced and/or how will existing data be re-used?
 - b. What data (for example the kinds, formats, and volumes) will be collected or produced?
2. **Documentation and data quality**
 - a. What metadata and documentation (for example the methodology of data collection and way of organising data) will accompany the data?
 - b. What data quality control measures will be used?
3. **Storage and backup during the research process**
 - a. How will data and metadata be stored and backed up during the research process?
 - b. How will data security and protection of sensitive data be taken care of during the research?
4. **Legal and ethical requirements, codes of conduct**
 - a. If personal data are processed, how will compliance with legislation on personal data and on data security be ensured?
 - b. How will other legal issues, such as intellectual property right and ownership, be managed? What legislation is applicable?
 - c. How will possible ethical issues be taken into account, and codes of conduct followed?
5. **Data sharing and long-term preservation**
 - a. How and when will data be shared? Are there possible restrictions to data sharing or embargo reasons?
 - b. How will data for preservation be selected, and where will data be preserved long-term (for example a data repository or archive)?
 - c. What methods or software tools will be needed to access and use the data?
 - d. How will the application of a unique and persistent identifier (such as a Digital Object Identifier (DOI)) to each data set be ensured?
6. **Data management responsibilities and resources**
 - a. Who (for example role, position, and institution) will be responsible for data management (i.e. the data steward)?
 - b. What resources (for example financial and time) will be dedicated to data management and ensuring that data will be FAIR (Findable, Accessible, Interoperable, Re-usable)?



The screenshot shows the homepage of the DMPonline website. At the top, there is a navigation bar with the DMPonline logo and links for Home, Public DMPs, Funder requirements, and Help. Below the navigation bar, a 'Welcome' message states that DMPonline helps users create, review, and share data management plans that meet institutional and funder requirements, and is provided by the Digital Curation Centre (DCC). It invites users to join a growing international community that has adopted DMPonline. Four statistics are displayed in a grid: 17,622 Users (with a person icon), 203 Organisations (with a classical building icon), 23,083 Plans (with a document icon), and 89 Countries (with a globe icon). At the bottom, a paragraph explains that some funders mandate the use of DMPonline, while others point to it as a useful option. It mentions that users can download funder templates without logging in, but the tool provides tailored guidance and example answers from the DCC and many research organisations, encouraging users to sign up for an account and try it out.





← → ↺ 🔒 dmponline.dcc.ac.uk

DMPONLINE Home Public DMPs Funder requirements Help

Welcome

DMPonline helps you to create, review, and share data management plans that meet institutional and funder requirements. It is provided by the Digital Curation Centre (DCC).

Join the growing international community that have adopted DMPonline:

- 
17,622 Users
- 
203 Organisations
- 
23,083 Plans
- 
89 Countries

Some funders mandate the use of DMPonline, while others point to it as a useful option. You can [download funder templates](#) without logging in, but the tool provides tailored guidance and example answers from the DCC and many research organisations. Why not sign up for an account and try it out?



Our resources

29.01.2020



Implementing Research Data Management Policies Across Europe: Experiences from Science Europe Member Organisations

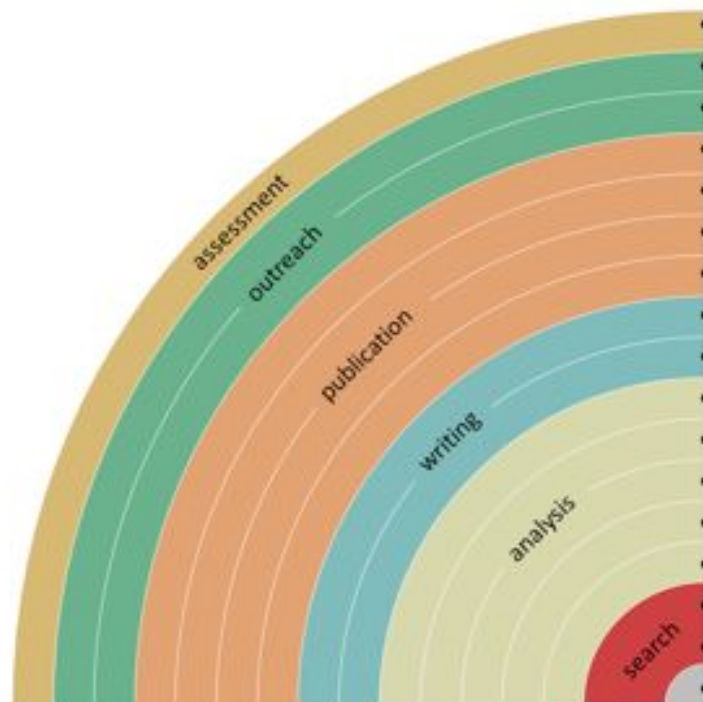
Does your organisation want to develop requirements for data management plans (DMPs) or update existing ones?

Take a look at our latest publication to find out how to do so in three steps.

On 29 January 2019, Science Europe released its 'Practical Guide to the International Alignment of Research Data Management' (the RDM Guide) which has been taken up by several SE MOs.

Based on their experiences, this publication features their approaches in developing and implementing DMP requirements. It explores challenges met by researchers, their home institutions and the funding organisations, both during and after implementation of the new policies, and provides examples on how researchers can be supported in their RDM efforts.

You can make your workflow more open by ...



- adding alternative evaluation, e.g. with altmetrics
- communicating through social media, e.g. Twitter
- sharing posters & presentations, e.g. at FigShare
- using open licenses, e.g. CC0 or CC-BY
- publishing open access, 'green' or 'gold'
- using open peer review, e.g. at journals or PubPeer
- sharing preprints, e.g. at OSF, arXiv or bioRxiv
- using actionable formats, e.g. with Jupyter or CoCalc
- open XML-drafting, e.g. at Overleaf or Authorea
- sharing protocols & workfl., e.g. at Protocols.io
- sharing notebooks, e.g. at OpenNotebookScience
- sharing code, e.g. at GitHub with GNU/MIT license
- sharing data, e.g. at Dryad, Zenodo or Dataverse
- pre-registering, e.g. at OSF or AsPredicted
- commenting openly, e.g. with Hypothes.is
- using shared reference libraries, e.g. with Zotero
- sharing (grant) proposals, e.g. at RIO



Possible learning goals for a training on research data management

Researcher-> Make DMP by using an online DMP tool

Service provider -> Select DMP tool / adjust online DMP tool


Policy maker -> Define policy on the application of a DMP

**Thank you
for your attention!**

Questions?



EOSC-hub

 eosc-hub.eu  [@EOSC_eu](https://twitter.com/EOSC_eu)



This material by Parties of the EOSC-hub Consortium is licensed under a Creative Commons Attribution 4.0 International License.

Topic 4 Cataloguing research output at the national level

Cataloguing Research Output at the National Level

The Dutch NARCIS as a case study

By Cees Hof (DANS)

Introduction

- Cataloguing in relation to the FAIR principals
- Cataloguing landscape analysis (small)

Overview of available possibilities

- Classification systems
- Training subjects

Practical training experiences

- Using the NARCIS case study



- We will focus on cataloguing at the national level
- Domain specific catalogues will be considered as sources for a national catalogue only
- Focus on metadata
- Cataloguing of all research output, open and restricted access
- Publications, data, software, researchers, others...
- For trainers that have a national catalogue or not

For train the trainers we are pionering!



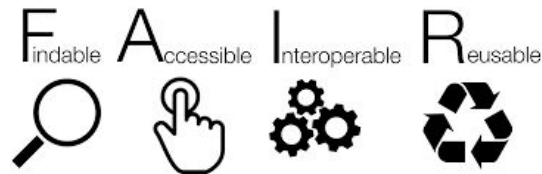
<https://dictionary.casrai.org/>



Use the CASRAI site for the consequent use of your terminology

Catalogue – A type of collection that describes, and points to features of another collection.

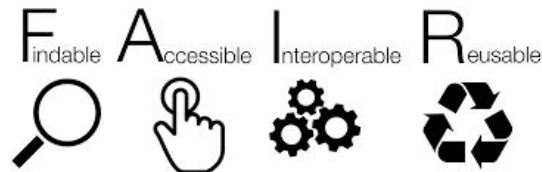
Registry – A database containing information about trusted repositories that are provided by the repository managers and are useful for human and machine users.



The use of catalogues is pivotal to monitor research output and increases the **Findability** of research output.

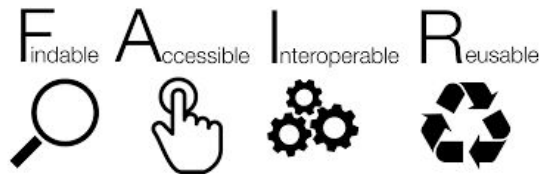


....as the information forest is getting more and more dense....



National catalogues, characteristics:

- Provide an **overview** of research output at the national level
- Present **metadata** but with a link to the source of the original data
- **Aggregate** metadata from distributed sources
- Are based on **standardised** metadata and metadata exchange protocols
- Usually focus on **publications**, increasingly also on **data** and other output such as **software** or **grey literature**

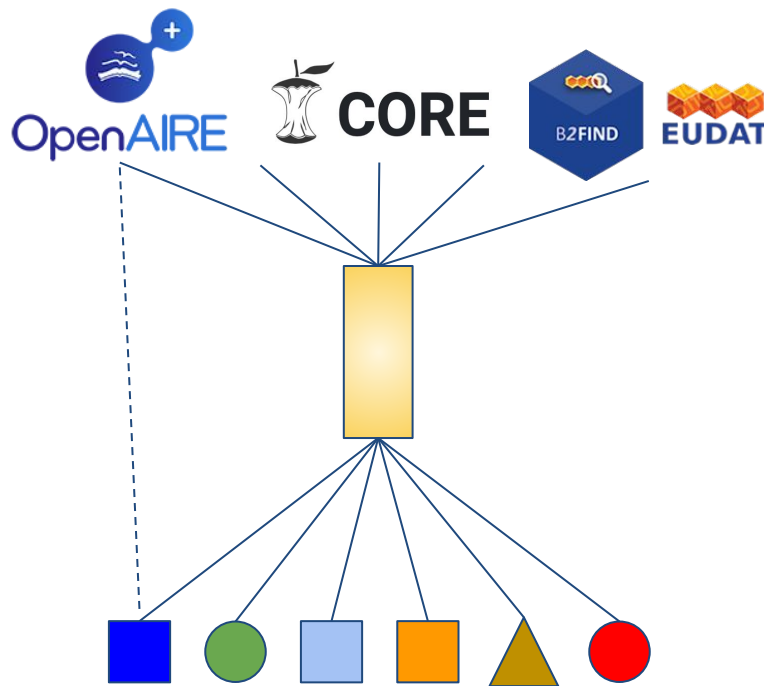


National catalogues, why important?

- Allow **statistics** over research output
- Provide **credits** to researchers and research groups
- Allow policy makers to **monitor** the results of funding programmes
- Can be used to **monitor open science** efforts
- Generate **new research opportunities** such as meta-analysis
- Provide metadata to other networks and initiatives

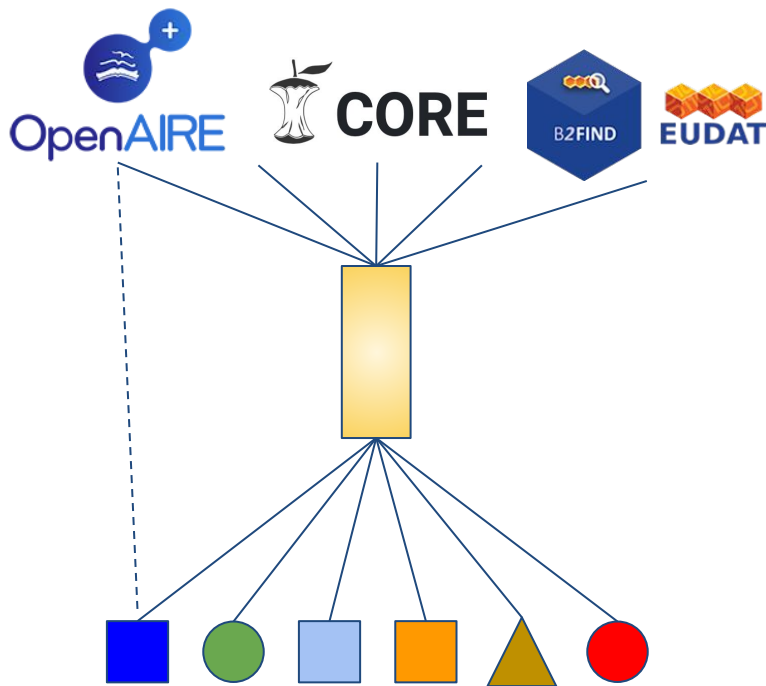
Increased
findability:

**By connecting
catalogue
systems you
multiply the
findability of
your research
output!**



Increased
findability:

**By connecting
catalogue
systems you
multiply the
findability of
your research
output!**



**International
Aggregators**

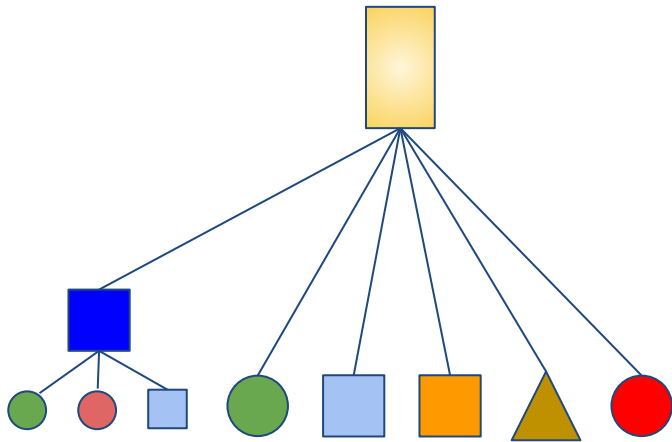


**Your national
Catalogue**

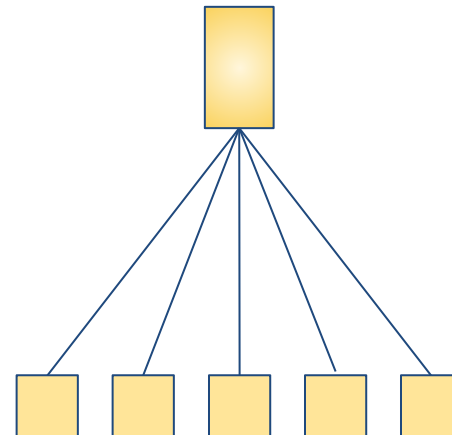


**Local
Repositories**

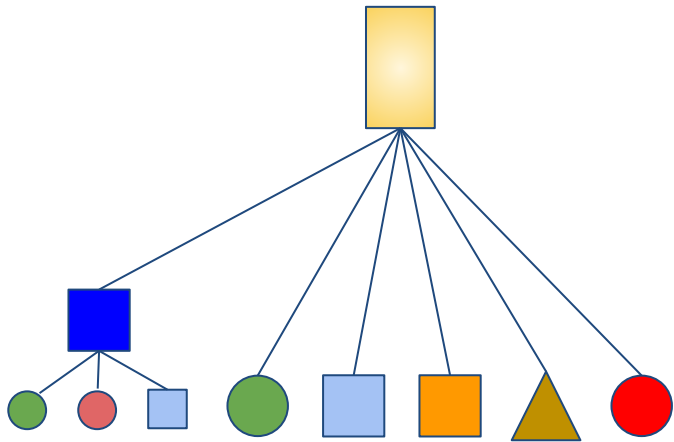
Systems: **Distributed open systems,
based on common metadata
standards and exchange
protocols**



**Distributed systems, based
on a common software
package/network.**

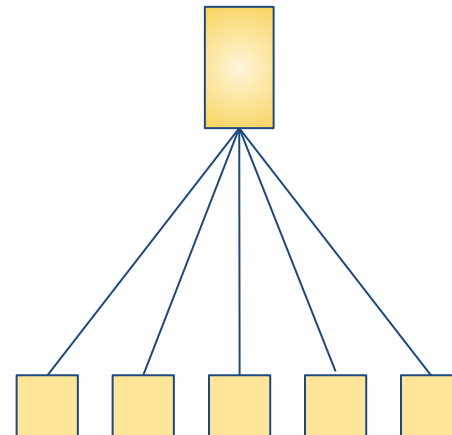


Systems: Distributed open systems,
based on common metadata
standards and exchange
protocols



**And there are
obviously many in
between and
cross-over
approaches...**

**Distributed systems, based
on a common software
package/network.**



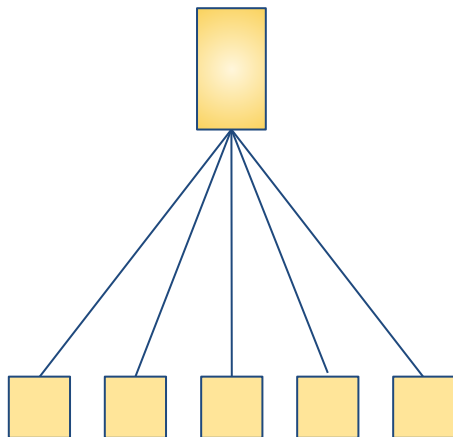
For example:



CRISTIN

Current Research Information System in Norway

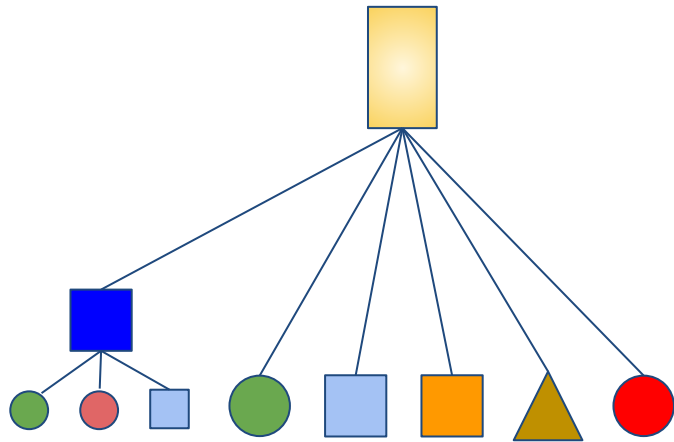
<https://www.cristin.no/english/>



- Common centralised system
- User accounts
- (self) Registering and reporting of research activities and results
- Member institutions in the research institute sector, the higher education sector and health trusts

For example: **NARCIS**

<https://www.narcis.nl/>




- NARCIS is build on a **Fedora** (DuraSpace) open source repository system.
- **Open Archives Initiative - Protocol for Metadata Harvesting** (OAI-PMH), main exchange protocol.
- Next to some Dutch metadata standards, **DublinCore** and **DataCite** are the most important standards.
- A national **classification system** for academic disciplines is being used.
- Build on top of 20+ **Current Research Information System(s)** (CRIS).
- Harvesting form **stand-alone repositories** and **databases**.

**Trainers
TIP!**

Use OpenAIRE to
see what is going
on in this
landscape...

<https://explore.openaire.eu>

 OpenAIRE EXPLORE			
SEARCH DEPOSIT LINK CONTENT PROVIDERS			
PUBLICATIONS	Funder	Project	Publication Date
RESEARCH DATA	European Commission (11,417)	ASIA (1,143)	2015 (1,442,748)
SOFTWARE	Tara Expeditions Fou... (550)	ERA-CLIM (971)	2018 (1,289,847)
OTHER RESEARCH PRODUCTS	Swiss National Scien... (515)	CARBOCHANGE (834)	2017 (1,121,858)
PROJECTS	National Institutes ... (238)	EMSODEV (733)	2019 (1,030,821)
CONTENT PROVIDERS	Wellcome Trust (98)	HERMIONE (689)	2016 (959,071)
ORGANIZATIONS	View more	View more	View more
	Access Mode	Type	Language
	not available (8,743,111)	Dataset (8,189,531)	English (2,294,838)
	Open Access (1,161,253)	Image (1,429,910)	en (1,122,826)
	Restricted (18,225)	Clinical Trial (147,456)	de (499,381)
	Embargo (13,114)	Audiovisual (135,451)	en-us (166,918)
	Closed Access (6,757)	Sound (31,777)	nl (91,291)
	View more	View more	View more
	Community	Content Provider	Collected From
	FET H2020 (118)	Unknown Repository (7,532,419)	Datacite (7,531,965)
	EGI Federation (42)	figshare (1,441,855)	figshare (1,441,855)
	FET FP7 (34)	figshare (1,441,855)	Omics Discovery I... (415,552)
	Research Data Alliance (4)	ENA (European Nuc... (378,881)	OpenTrials (138,888)



What to train when it comes to cataloguing research output?

Training topics: we stay out of technology.... but focus on data managers and data stewards:

- **Mapping Metadata**
- **Rich & Enriching Metadata**
- **Metadata awareness**
- How to use **Classification Systems** for academic disciplines
- Using and implementing local **Current Research Information System(s)** (CRIS) (not part of this webinar)

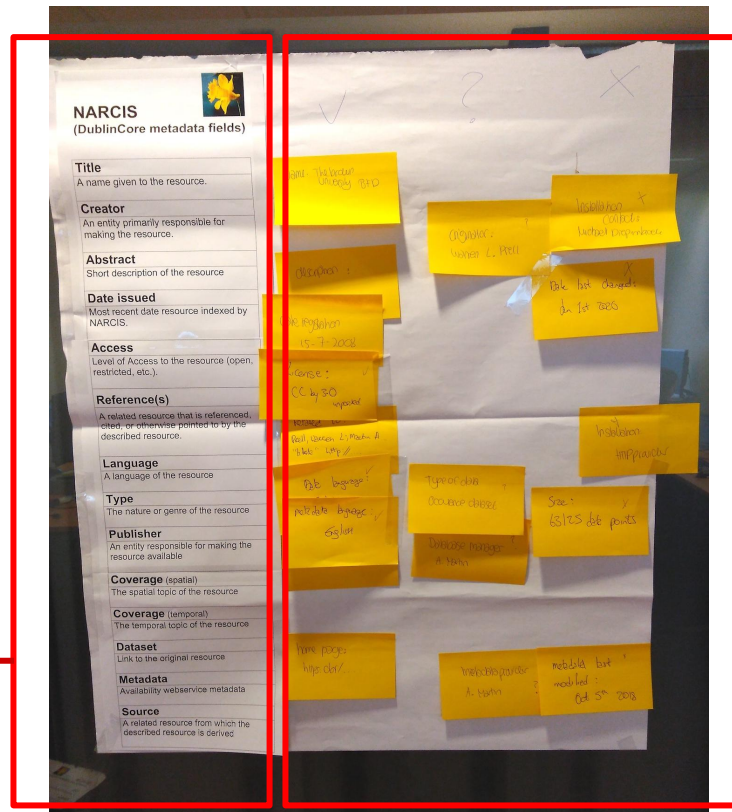


- **Mapping Metadata**

How well do your “local” metadata fit the national catalogue metadata?

(Or generic metadata standards if there is no catalogue yet...)

NARCIS catalogue
DublinCore metadata



Dutch
BioBank
metadata



- **Rich & Enriching Metadata**

If your metadata become part of a larger system, and different portals, how to increase their usability and visibility?

For example:

DublinCore "coverage" = Texel

Much better also:

DublinCore "coverage" =

[4.691162](#), [52.981723](#), [4.913635](#), [53.192047](#) (Bounding Box)





- How to use **Classification Systems** for academic disciplines?

D20000	Life sciences, medicine and health care	
	D21000	Life sciences
	D21100	Bioinformatics, biomathematics
	D21200	Biophysics, clinical physics
	D21300	Biochemistry, molecular biology
	D21400	Genetics
	D21500	Histology, cell biology
	D21600	Anatomy, morphology
	D21700	Physiology
	D21800	Immunology, serology
	D21900	Epidemiology and medical statistics
	D22000	Biology
	D22100	Microbiology
	D22200	Biogeography, taxonomy
	D22300	Animal ethology, animal psychology
	D22400	Ecology
	D22500	Botany
	D22600	Zoology
	D22700	Toxicology (plants, invertebrates)

- Which classifications to use?
- What are the preferences of certain communities
- The effects of using classifications on metrics and statistics of national catalogues
- The ins and outs of using codes
-

<https://www.narcis.nl/classification/Language/en>



- **Metadata awareness....**

Catalogues showcase your data through metadata...

- Training on the use of vocabularies
- Training on language issues
- Training around GDPR issues in metadata
- Writing skills for research descriptions
-




**Thank you
for your attention!**

Questions?



EOSC-hub

 eosc-hub.eu  [@EOSC_eu](https://twitter.com/EOSC_eu)



This material by Parties of the EOsc-hub Consortium is licensed under a Creative Commons Attribution 4.0 International License.

Please fill in the evaluation form at:


<https://forms.gle/mkzWwZxx8DRdfidi9>

**Thank you
for your attention!**

Questions?



EOSC-hub

 eosc-hub.eu  [@EOSC_eu](https://twitter.com/EOSC_eu)



This material by Parties of the EOsc-hub Consortium is licensed under a Creative Commons Attribution 4.0 International License.