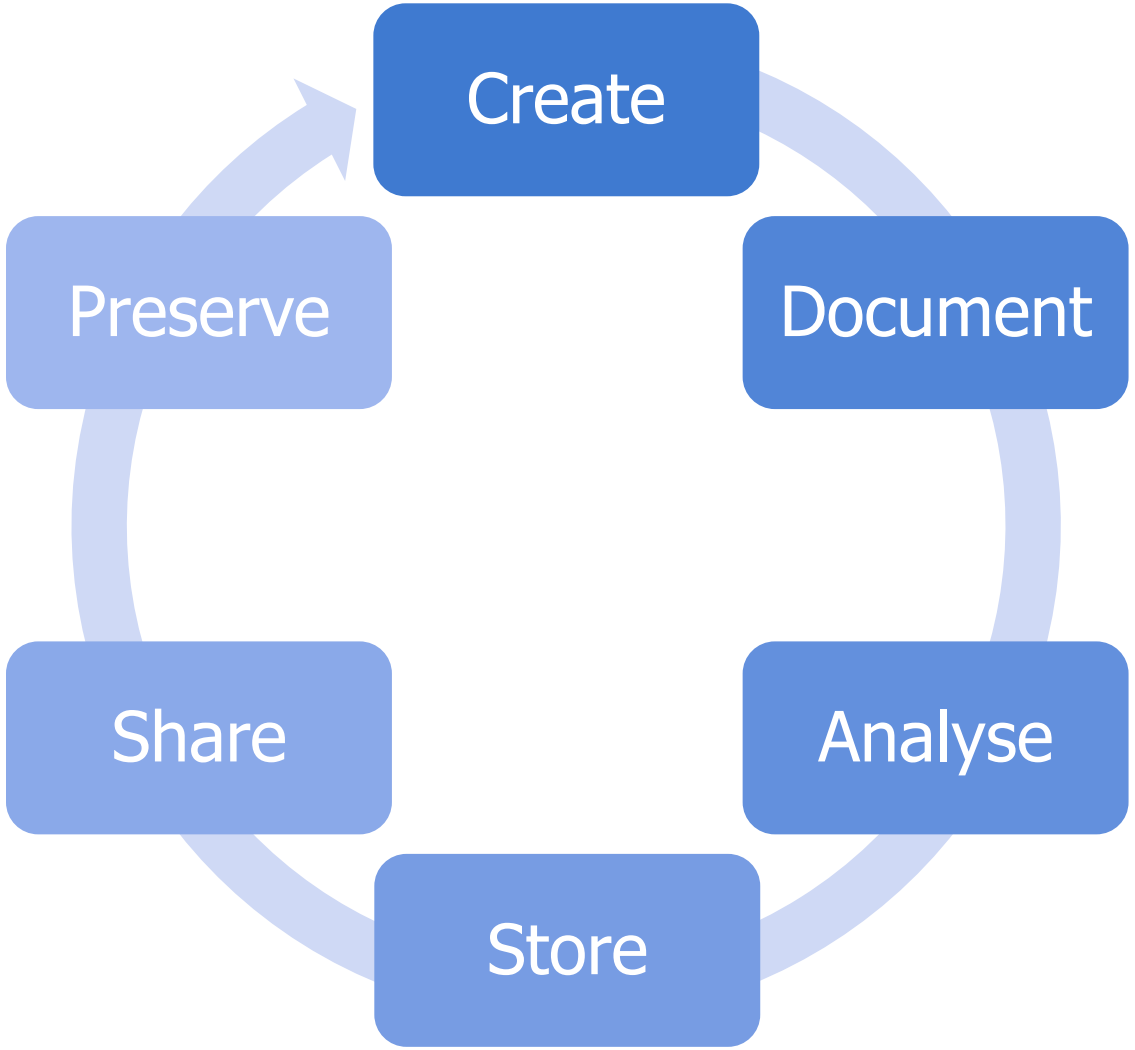
Data Curation Lifecycle

S. Venkataraman, Training Officer, OpenAIRE

Third OpenAIRE Train-the-Trainers Open Science Bootcamp

24th May 2022 (Virtual)

THE RESEARCH DATA CURATION LIFECYCLE

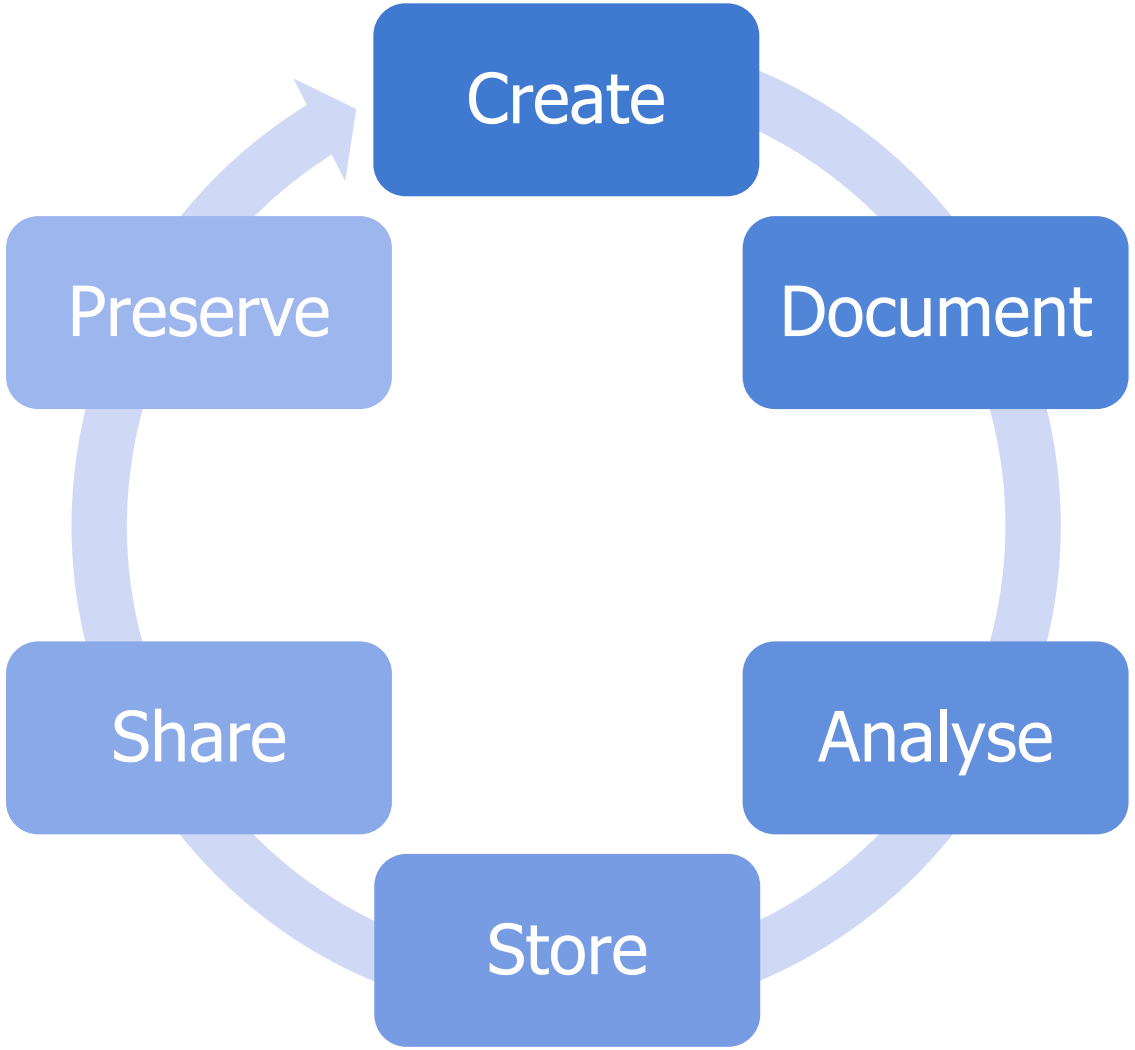


Adapted from DCC.

WHAT IS RESEARCH DATA MANAGEMENT?

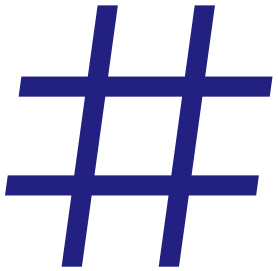
"the active management and appraisal of data over the lifecycle of scholarly and scientific interest"

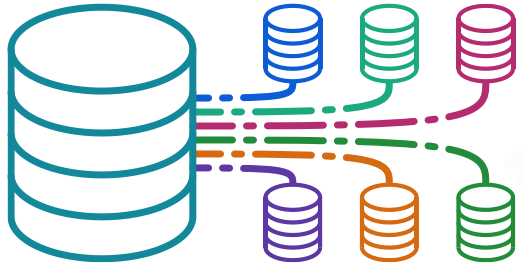
Data management is part of good research practice.



Adapted from DCC.

Good RDM involves a series of measures…





Enforcing existing policies…

This policy for managing research data was approved by the University Court on 16 May, 2011.

The University adopts the following policy on Research Data Management. It is acknowledged that this is an aspirational policy, and that implementation will take some years.

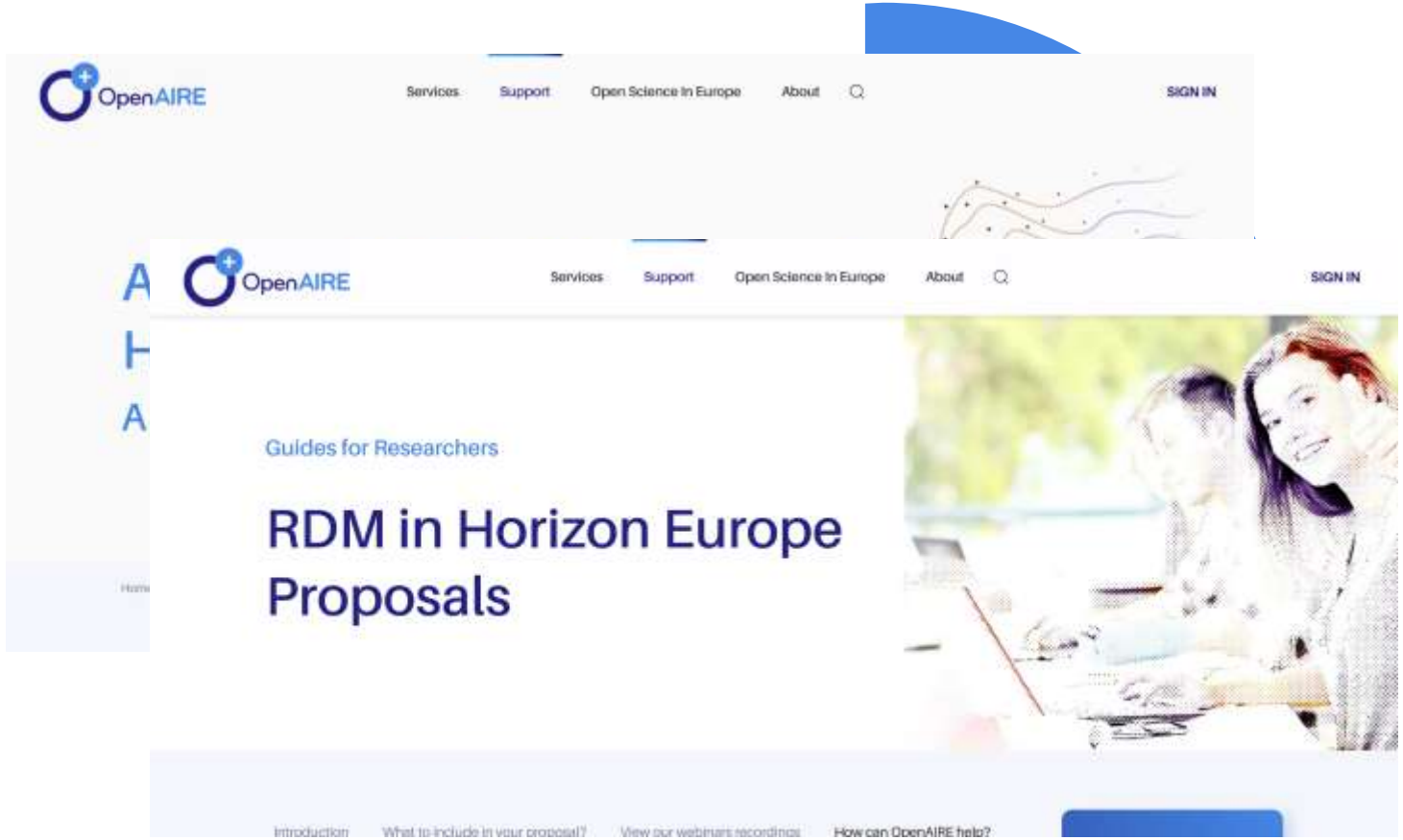
1. Research data will be managed to the highest standards throughout the research data lifecycle as part of the University's commitment to research excellence.
2. Responsibility for research data management through a sound research data management plan during any research project or programme lies primarily with Principal Investigators (PIs).
3. All new research proposals [from date of adoption] must include research data management plans or protocols that explicitly address data capture, management, integrity, confidentiality, retention, sharing and publication.
4. The University will provide training, support, advice and where appropriate guidelines and templates for the research data management and research data management plans.
5. The University will provide mechanisms and services for storage, backup, registration, deposit and retention of research data assets in support of current and future access, during and after completion of research projects.
6. Any data which is retained elsewhere, for example in an international data service or domain repository should be registered with the University.
7. Research data management plans must ensure that research data are available for access and re-use where appropriate and under appropriate safeguards.
8. The legitimate interests of the subjects of research data must be protected.
9. Research data of future historical interest, and all research data that represent records of the University, including data that substantiate research findings, will be offered and assessed for deposit and retention in an appropriate national or international data service or domain repository, or a University repository.
10. Exclusive rights to reuse or publish research data should not be handed over to commercial publishers or agents without retaining the rights to make the data openly available for re-use, unless this is a condition of funding.

Related Information

- University Research Data Service
- University Research Data Roadmap

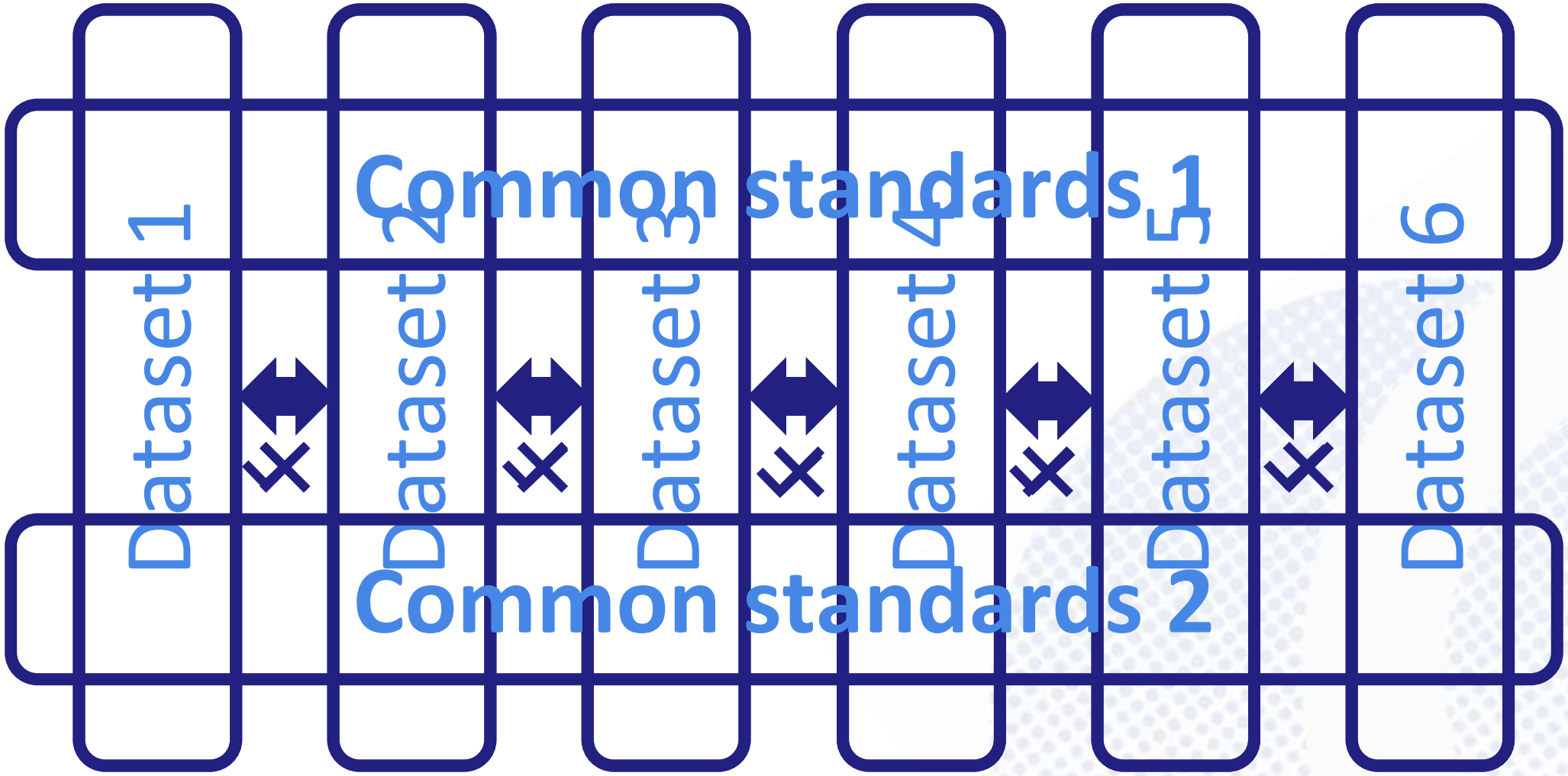
https://www.ed.ac.uk/information-services/about/policies-and-regulations/research-data-policy

OPENAIRE GUIDES



https://www.openaire.eu/rdm-handbook
https://www.openaire.eu/rdm-in-horizon-europe-proposals

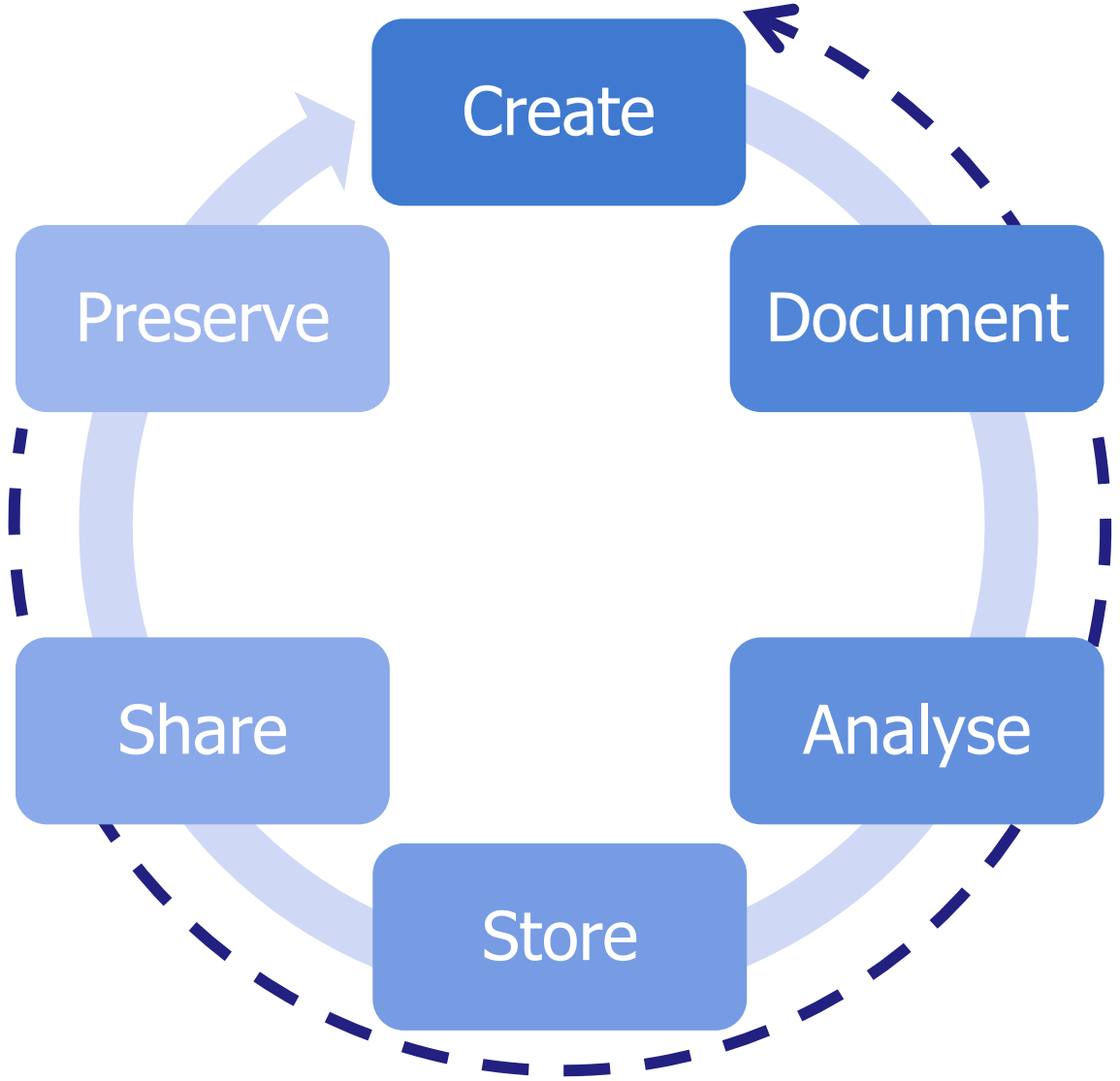
Open vs FAIR


Common standards 1
Dataset 1
Dataset 2
Dataset 3
Dataset 4
Dataset 5
Dataset 6
Common standards 2

Data silos

OPEN SCIENCE

- Change the typical lifecycle
- Publish earlier and release more
- Papers + Data + Methods + Code…
- Support reproducibility



Adapted from DCC.

…as opposed to…

Findable **A**ccessible **I**nteroperable **R**eusable

- Metadata
- PIDs
- Repositories

- Metadata
- Open file formats and software

- Metadata
- Ontologies
- Repositories

- Metadata
- Licences

What FAIR means: 15 principles

Findable:

F1. (meta)data are assigned a globally unique and persistent identifier;

F2. data are described with rich metadata;

F3. metadata clearly and explicitly include the identifier of the data it describes;

F4. (meta)data are registered or indexed in a searchable resource;

Accessible:

A1. (meta)data are retrievable by their identifier using a standardized communications protocol;

A1.1 the protocol is open, free, and universally implementable;

A1.2. the protocol allows for an authentication and authorization procedure, where necessary;

A2. metadata are accessible, even when the data are no longer available;

Interoperable:

I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

I2. (meta)data use vocabularies that follow FAIR principles;

I3. (meta)data include qualified references to other (meta)data;

Reusable:

R1. meta(data) are richly described with a plurality of accurate and relevant attributes;

R1.1. (meta)data are released with a clear and accessible data usage license;

R1.2. (meta)data are associated with detailed provenance;

R1.3. (meta)data meet domain-relevant community standards;

doi: 10.1038/sdata.2016.18

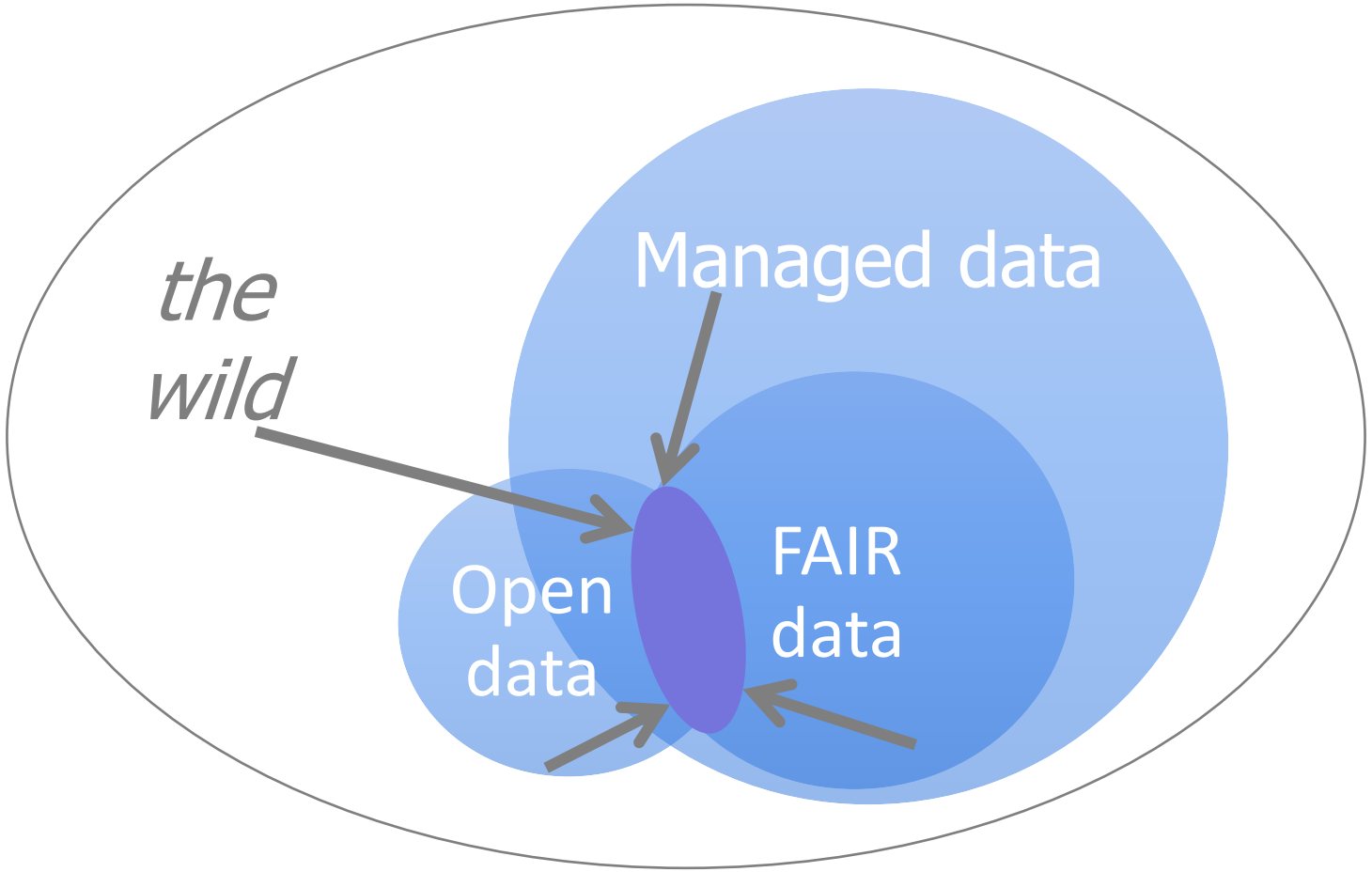
Comprehensive descriptions can be found at https://www.go-fair.org/fair-principles/

But what about awareness levels?

“We know that implementing open science practices can be beneficial for individual researchers, and as might be hoped, the flurry of open science activity has resulted in researcher awareness of the FAIR data principles steadily increasing. In the State of Open Data 2018 we found that 15% of the respondents were familiar with FAIR, increasing to 18% and 20% in 2019 and 2020 respectively. However, familiarity does not necessarily translate to understanding and in our State of Open Data surveys we also observed a rise in the proportion of researchers who had heard of FAIR but did not consider themselves familiar with the principles; 25% in 2018, 28% in 2019 and 31% in 2020. This demonstrates a persistent gap in researchers' understanding of how to implement the FAIR principles for their own work, which can mean that even when researchers do share their data it’s not necessarily in a usable form.”

https://www.springernature.com/gp/advancing-discovery/springboard/blog/blogposts-open-research/how-fair-is-driving-open-research-forwards/19189878

Increasing that which is FAIR & open



Adapted from DCC.

FAIR ≠ OPEN

as open as possible, as closed as necessary

Image: ‘Balancing rocks’ by Viewminder CC-BY-SA-ND www.flickr.com/photos/light_seeker/7780857224

Adapted from DCC.

FAIR ISN'T THE ONLY THING…

EC POLICY

EUROPEAN OPEN SCIENCE CLOUD

THANKS

Location
www.openaire.eu

Email
s.venkataraman@openaire.eu

Twitter
@openaire_eu
#OpenAIRE_OSBootcamp