



Data  
Schools

# RDM & Open Science: fundamentals

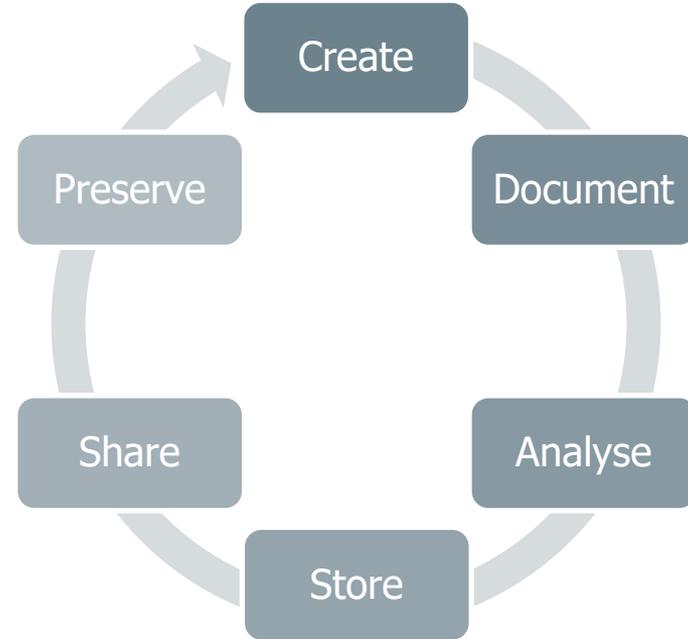
*S. Venkataraman, OpenAIRE*



# What is Research Data Management?

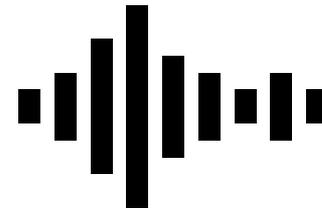
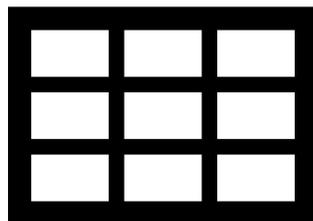
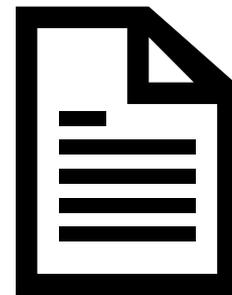
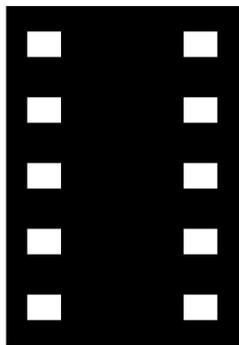
“the active management  
and appraisal of data over  
the lifecycle of scholarly and  
scientific interest”

**Data management is part  
of good research practice.**



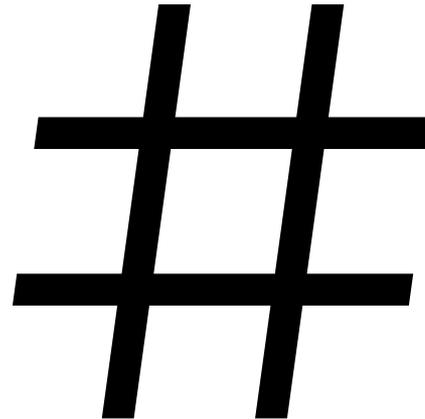
# File formats

- Avoid **proprietary** and **lossy** formats
- Make sure **accessibility** ensured in long term preservation



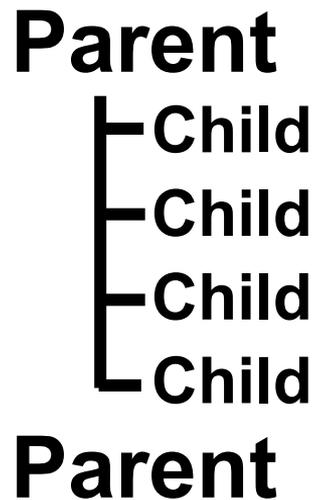
# Metadata

- “data about data”
- Part of good documentation
- **Minimum information**
- Use common standards where possible to allow **interoperability**
- Look for standards here: <http://rd-alliance.github.io/metadata-directory/>;  
<https://rdamsc.dcc.ac.uk/>



# Ontologies

- Part of metadata
- **Structured controlled vocabularies**
- Use common ontologies where possible
- Allows easier cross comparisons to be done computationally -> **interoperability**
- Try [BARTOC catalogue](#) to possibly find suitable ontology for you



# Licences

- Provide information to any potential data **reuser** of their rights
- Ensures clarity
- Creative commons (CC) licensing most used in research

CREATIVE COMMONS LICENSES		COPY & PUBLISH	ATTRIBUTION REQUIRED	COMMERCIAL USE	MODIFY & ADAPT	CHANGE LICENSE
	PUBLIC DOMAIN	✓	✗	✓	✓	✓
	CC BY	✓	✓	✗	✓	✓
	CC BY-SA	✓	✓	✗	✗	✗
	CC BY-ND	✓	✓	✗	✗	✗
	CC BY-NC	✓	✓	✗	✓	✓
	CC BY-NC-SA	✓	✓	✗	✓	✗
	CC BY-NC-ND	✓	✓	✗	✗	✗

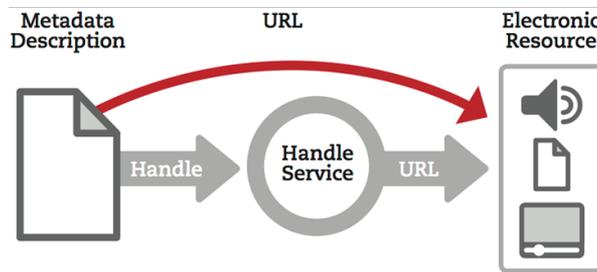
  

				
You can redistribute (copy, publish, display, communicate, etc.)	You have to attribute the original work	You can use the work commercially	You can modify and adapt the original work	You can choose license type for your adaptations of the work.

Part of [How To Attribute Creative Commons Photos](#) by Foter, licensed CC BY SA 3.0

# PIDs

- Unique, persistent identifiers can be used for different types of objects
- Ensures **disambiguation** and **findability**
- e.g. DOIs, ORCID, ISBNs



**Publication date:**  
November 24, 2017

**DOI:**  
DOI: [10.5281/zenodo.1065991](https://doi.org/10.5281/zenodo.1065991)

**Keyword(s):**  
FAIR, FAIRness, checklist, research data, Findable, Accessible, Interoperable, Reusable, PID, repository, DOI, metadata, licence, data sharing, research data management,

**Grants:**  
[European Commission](#)

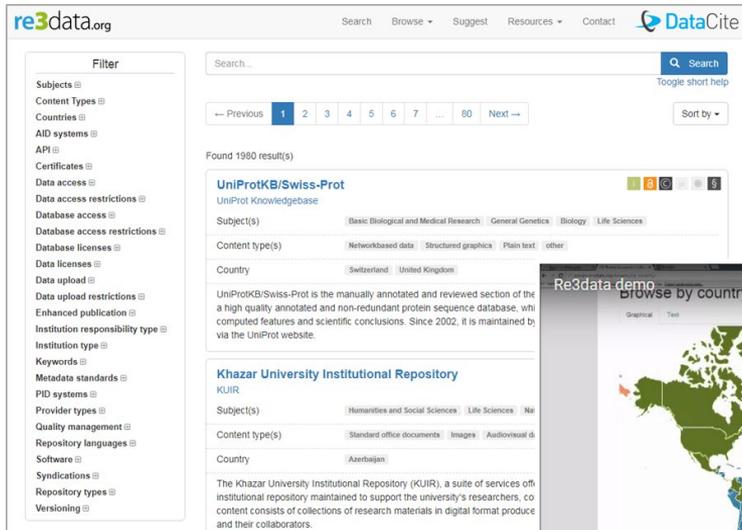
- EUDAT2020 - EUDAT2020 (654065)

**License (for files):**  
[Creative Commons Attribution 4.0](#)



# Repositories

- If possible, choose a suitable repository at the very start of a project
- Use [re3data.org](https://re3data.org) to search for repos
- Tip: many of the questions that arose previously may be answered by the choice of a repo in advance!



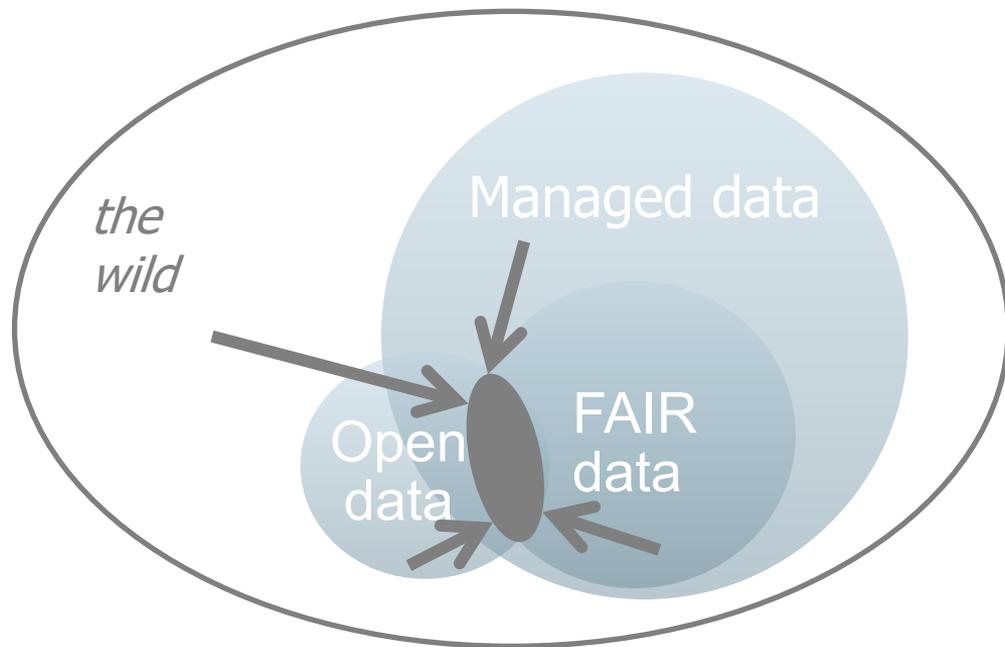
The screenshot shows the re3data.org search interface. On the left is a 'Filter' sidebar with categories like Subjects, Content Types, Countries, etc. The main search area shows results for 'UniProtKB/Swiss-Prot' and 'Khazar University Institutional Repository'. The UniProt result includes details about its subjects (Basic Biological and Medical Research, General Genetics, Biology, Life Sciences) and content types (Networkbased data, Structured graphics, Plain text, other). The Khazar result includes details about its subjects (Humanities and Social Sciences, Life Sciences) and content types (Standard office documents, Images, Audiovisual).





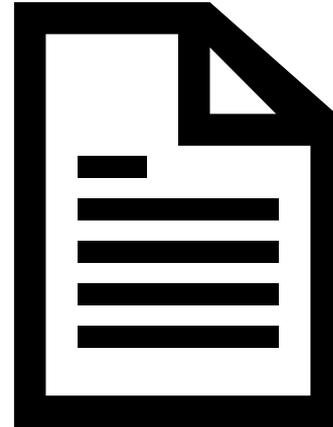
# Open Science

- Where possible, any publicly funded research outputs should be made publicly available
- FAIR ≠ Open
- “As open as possible, as closed as necessary”



# Data Management Plans (DMPs)

- A document written before the start of a project
- A “**living document**”
- A way of neatly tying together all the info discussed previously
- They are there to mainly help you!
- Try the DCC’s [DMPonline](#) tool to get you started



# Take home messages



- Increase **value**
- Increase **reproducibility**
- Increase **provenance**
- Increase **integrity**
- Increase **accountability**
- Reduce **risks**
- Reduce **costs**
- Reduce **fraud**



# Data Schools

[s.venkataraman@openaire.eu](mailto:s.venkataraman@openaire.eu)

[@OpenAIRE\\_eu](#)

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



OpenAIRE | NEXUS



# Homework

Look at the *Barriers to Open Science* document and the different arguments listed that one may face when there is resistance. Can you provide arguments that would change the minds of people that have these fears about being more open with their data?