



HAL
open science

Research data management, a chance for Open Science. Methods and tutorials to create a Data Management Plan (DMP)

Marie Puren, Charles Riondet

► To cite this version:

Marie Puren, Charles Riondet. Research data management, a chance for Open Science. Methods and tutorials to create a Data Management Plan (DMP). DARIAH's Humanities at Scale Winter School, Dariah; Humanities at Scale, Oct 2016, Prague, Czech Republic. hal-01416978

HAL Id: hal-01416978

<https://inria.hal.science/hal-01416978v1>

Submitted on 15 Dec 2016

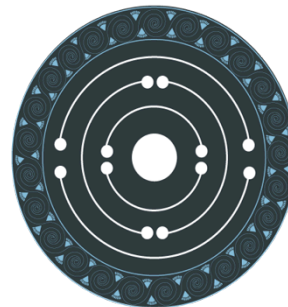
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Research data management, a chance for Open Science.

Methods and tutorials to create a Data Management Plan (DMP)

Marie Puren
Charles Riondet



PARTHENOS

Pooling Activities, Resources and Tools
for Heritage E-research Networking,
Optimization and Synergies

Introduction

“Effective management of data promises rewards **throughout and beyond the life of a research project**. [...] For the researcher, the perception of data as an instrument of research and new knowledge can be **transformational**. Well-managed data lead to **higher-quality research**, increased visibility and the consequent benefits of enhanced citation rates.”

[How to develop RDM Services](#), Digital Curation Center

Lesson Topics

- Open science
- The Research Data Management
- What are Data Management Plans (DMP)
- Components of a DMP
- Why prepare a DMP?
- How to make a DMP

Learning objectives

After completing this lesson, the participant will be able to:

- Define the Research Data Lifecycle
- Define a DMP
- Understand the importance of preparing a DMP
- Identify the key components of a DMP
- Create a DMP

What is research data?

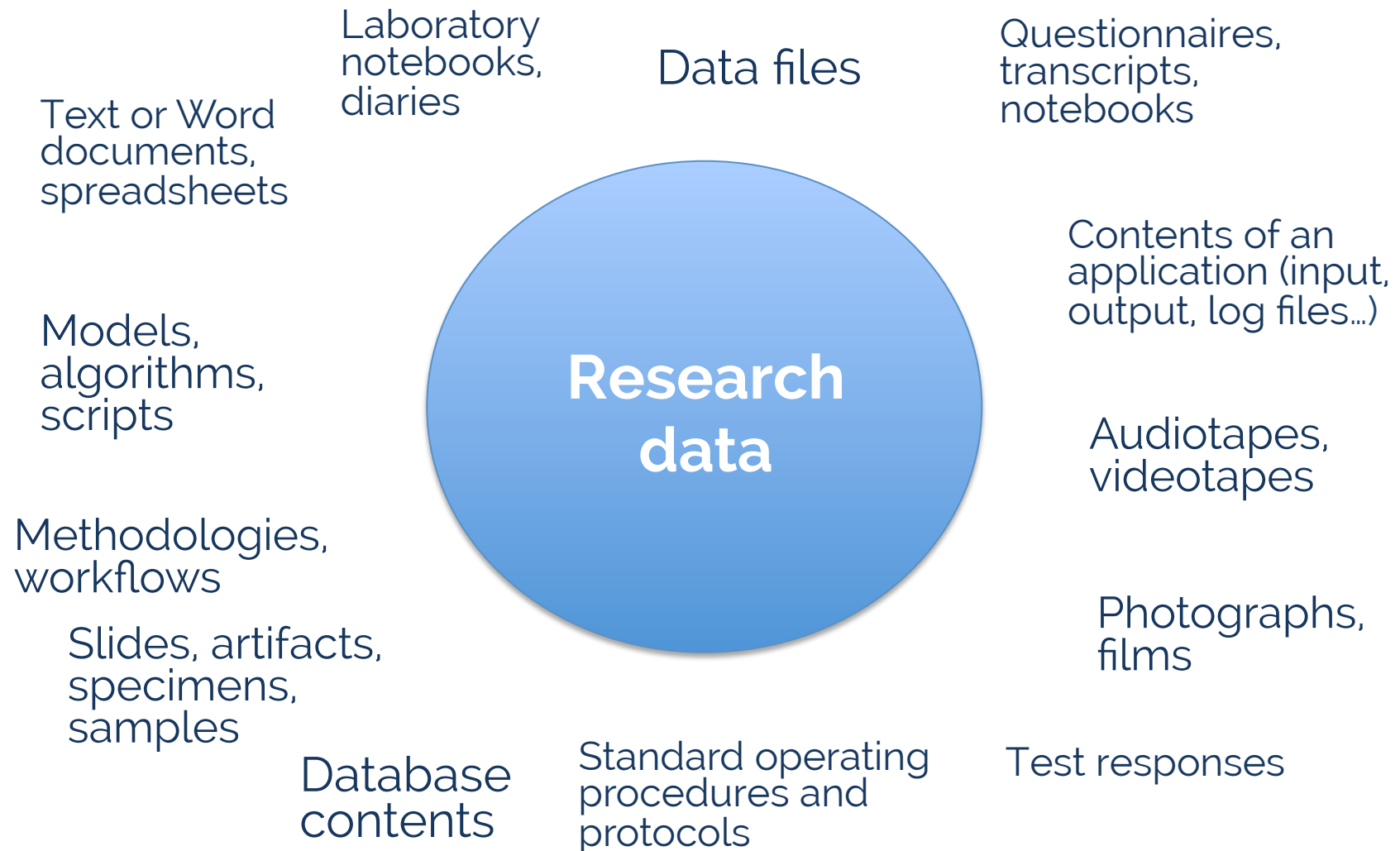
- “Data are distinct pieces of information, usually formatted in a special way.”
([BU Libraries](#))
- Defining “research data” is challenging.
- [BU Libraries](#): “Research data is data that is collected, observed, or created, for purposes of analysis to produce original research results.”

What is research data?

Research data can be:

- Observational
- Experimental
- Generated from test models (simulation)
- Derived or compiled (like text and data mining)
- Reference or canonical (for instance, gene sequence data banks)

What is research data?



What is research data?

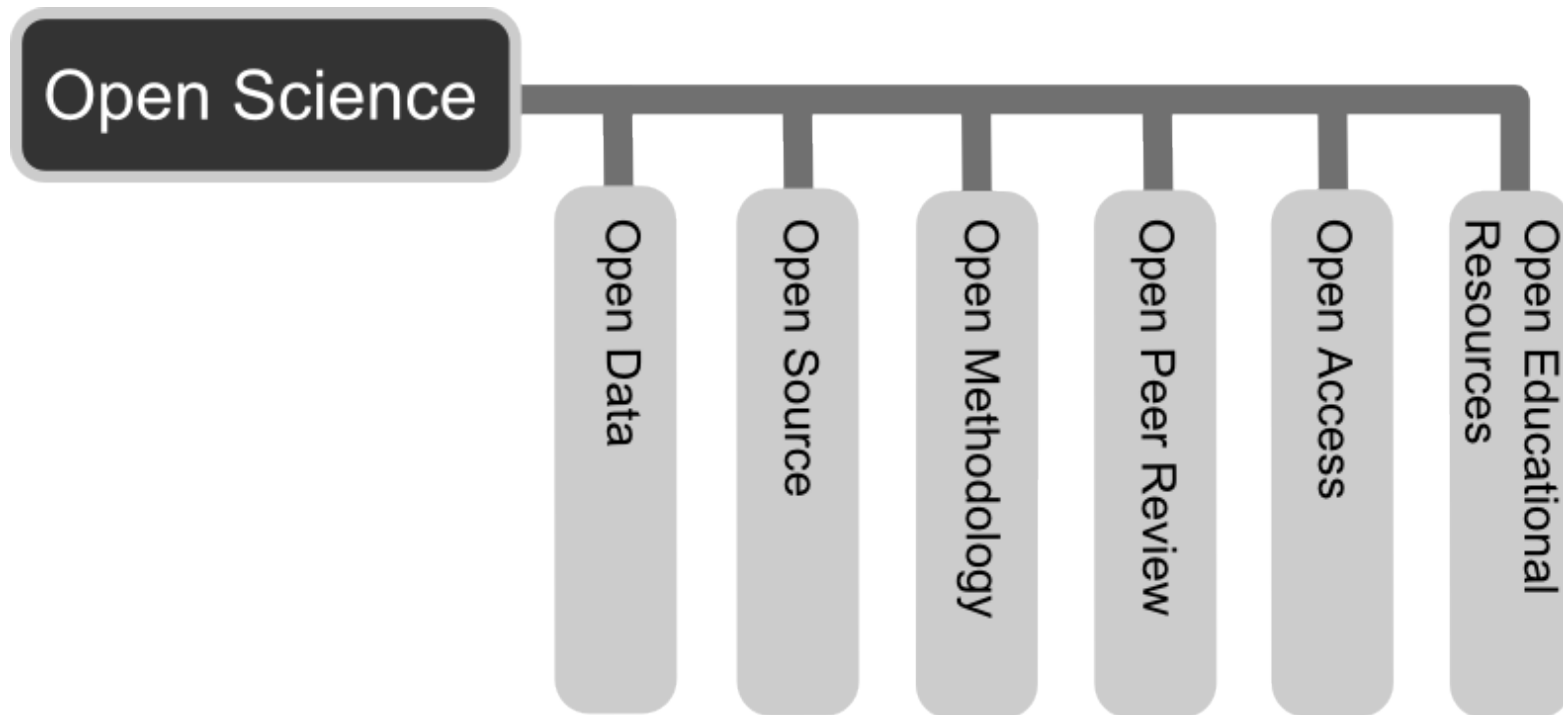
- Research data:
 - data that help to do research;
 - data that could be collected or created, then analysed;
 - data that come in multiple formats.
- Dataset: “might comprise a single element [...] [or] a collection of related elements.”

[\(Oxford Research Data Website\)](#)

A new model of openness for research data

- Availability and access
- Re-use and redistribution
- Universal participation

A new model of openness for research data



By [Andreas E. Neuhold](#), own work - based on "[The taxonomy tree](#)", FOSTER (Facilitate Open Science Training for European Research)

For more information on Open Science: Michael Nielsen, [Reinventing Discovery: The New Era of Networked Science](#), Princeton University Press, 2011.

A new model of openness for research data

Open = “Anyone can freely access, use, modify, and share for any purpose.”

Open Knowledge International, “[The Open definition](#)”

A new model of openness for research data

“It has become increasingly apparent that scientific data should be considered a product in much the same way journal articles or conference proceedings are [...].”

Felicia LeClere, [“Too Many Researchers Are Reluctant to Share Their Data”](#),
The Chronicle of Higher Education, 2010.

Supported by European and national initiatives

Horizon 2020 Research and Innovation Programme
The Pilot on Open Research Data (ORD Pilot)

“The ORD pilot applies primarily to the **data needed to validate the results** presented in scientific publications. **Other data** can also be provided by the beneficiaries **on a voluntary basis**, as stated in their Data Management Plans.”

[H2020 Programme Guidelines on FAIR Data Management in Horizon 2020](#), Version 3.0, 26 July 2016, p.3.

Supported by European and national initiatives

Extension of the ORD Pilot in July 2016

“The Commission has enabled access to and reuse of research data generated by Horizon 2020 projects through the Open Research Data Pilot (ORD Pilot). As stated in the 2017 work programme, the pilot is being extended to cover all thematic areas as described below. [...] By extending the pilot, open access becomes the default setting for research data generated in Horizon 2020.”

[H2020 Programme Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020](#), Version 3.0, 26 July 2016, p.8.

Supported by European and national initiatives

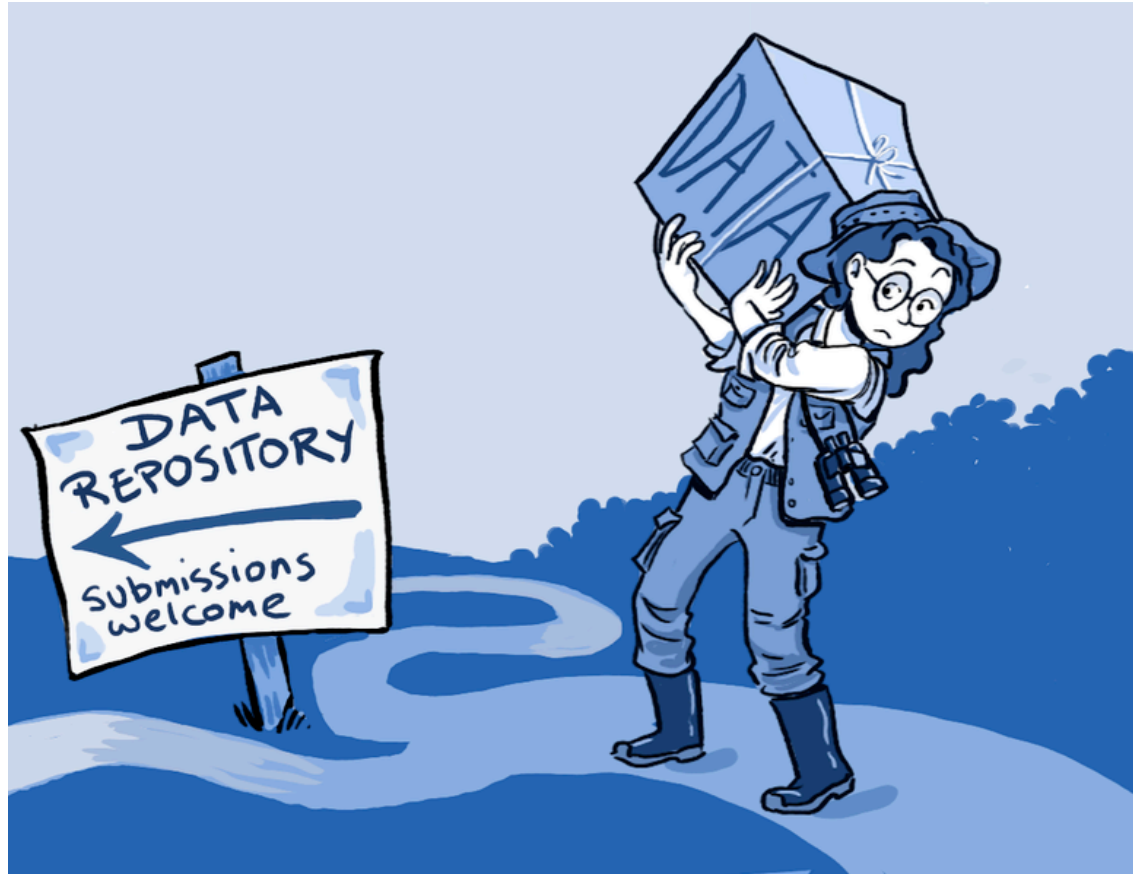
- United States:
 - *National Institutes of Health* (NIH) since 2003
 - *National Science Foundation* (NSF) since 2010
- United Kingdom:
 - *National Environment Research Council* (NERC), *Research Councils* and *Wellcome Trust* (2000-2010)
 - *Biotechnology and Biological Sciences Research Council* (BBSRC) *Data Sharing Policy* in 2007 (updated in 2010)
 - *JISC* (*Joint Information Systems Committee*) and *Digital Curation Center* (DCC) (2005)

Direct benefits for researchers

“Data sharing is a bit like going to the dentist. We can all agree that it is a good thing to do and intrinsic to good scientific practice. In reality, however, researchers tend to view data sharing with a mix of fear, contempt, and dread.”

Felicia LeClere, [“Too Many Researchers Are Reluctant to Share Their Data”](#),
The Chronicle of Higher Education, 2010.

Direct benefits for researchers



"To deposit or not to deposit, that is the question"

Roche DG, Lanfear R, Binning SA, Haff TM, Schwanz LE, et al. (2014)
"Troubleshooting Public Data Archiving: Suggestions to Increase Participation",
PLoS Biol 12(1): e1001779. [doi:10.1371/journal.pbio.1001779](https://doi.org/10.1371/journal.pbio.1001779)

Direct benefits for researchers

1. Fulfill requirements
2. Increase your research impact and visibility
3. Save time
4. Preserve your data
5. Ensure higher quality data

Direct benefits for researchers

6. Promote innovation and potential new data uses
7. Maximise transparency and accountability
8. Support Open Access
9. Help less rich institutions and countries to do research
10. Make good science

Direct benefits for funders

1. Maximise return on investment
2. Reduce the cost of duplicating data collection
3. Have access to great resources for education and training

Sharing data

How can I do that?



Managing your research data

Why manage data?

For yourself:

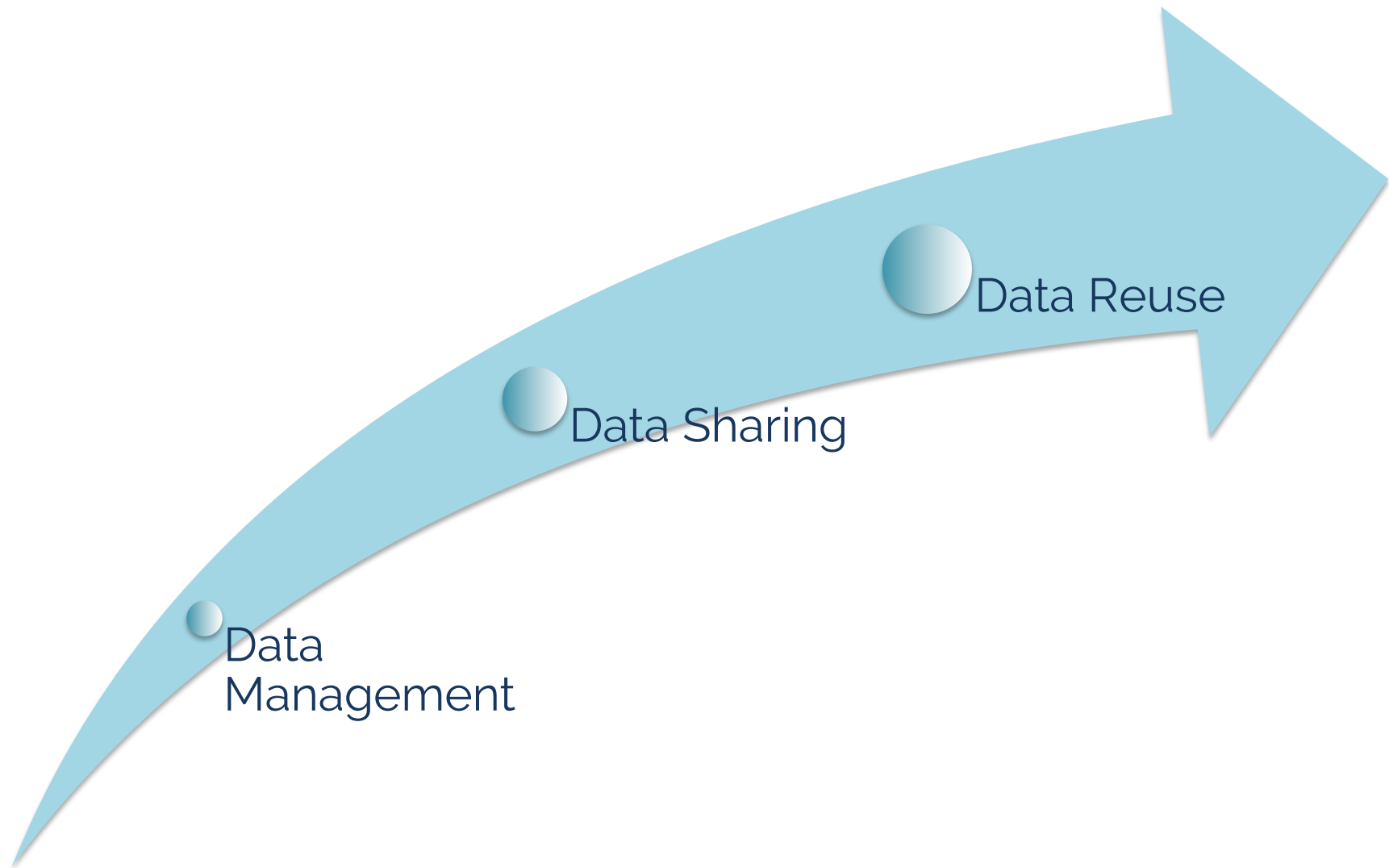
- Keep yourself organized: be able to find your files;
- Control the various versions of your data;
- Quality control your data more efficiently;
- Make backups to avoid data loss;
- Format your data for re-use (by yourself or others);
- Be prepared: document your data for your own recollection and re-use (by yourself or others).

Why manage data?

For funders:

- Maximize the effective use and value of data and information assets;
- Be assured that the quality of data is continually improved;
- Ensure appropriate use of data and information;
- Facilitate data sharing;
- Ensure sustainability and accessibility for re-use in science.

Why manage data?



Research data management (or RDM)

Definition

“Data management is integral to the process of conducting research.”

University of Leicester, [“When do you need to think about RDM”](#)

During a research project, and after the project is complete:

- Collect,
- Organise,
- Manage,
- Store,
- Back up,
- Preserve,
- Share your data.

Research data management (or RDM)

Definition

“Good research data management is not a goal in itself, but rather the key conduit leading to knowledge discovery and innovation, and to subsequent data and knowledge integration and reuse.”

[Guidelines on FAIR Data Management in Horizon 2020](#), Version 3.0, 26 July 2016, p.3.

Research data management (or RDM)

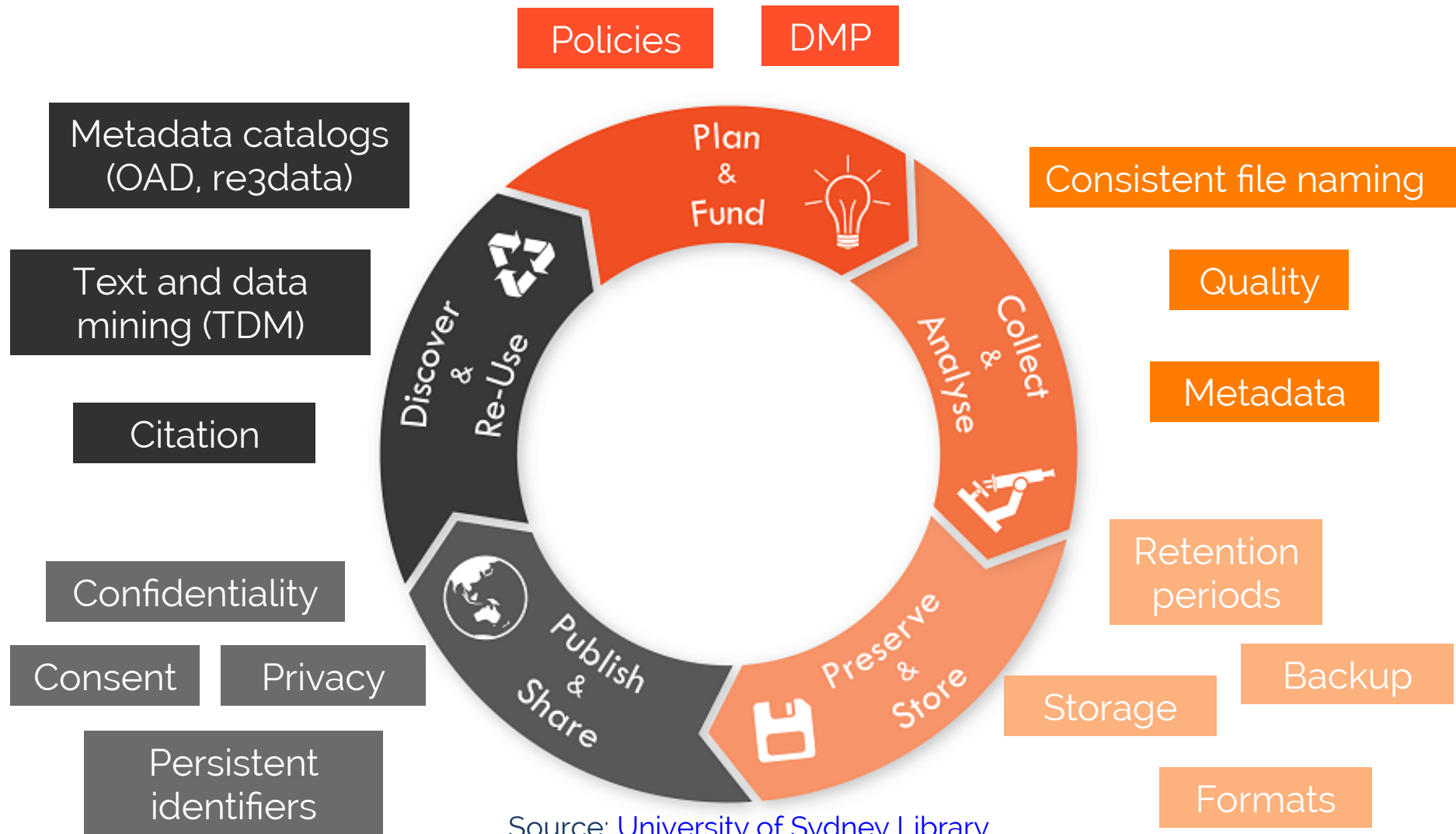
Definition

Research data management involves:

- Creating a **Data Management Plan** (or DMP);
- Storing research data;
- Depositing data in a long-term.

Research data management (or RDM)

Research data lifecycle



Source: [University of Sydney Library](#)

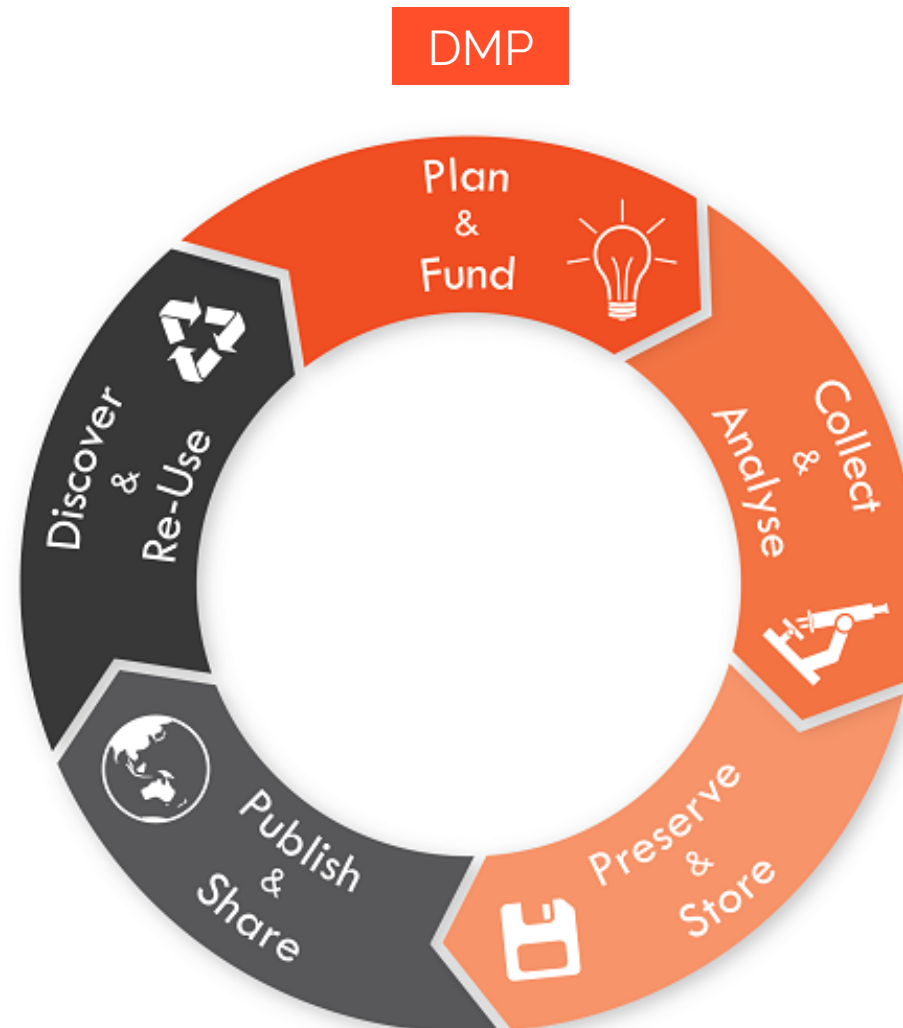
Research data management (or RDM)

Developing RDM services



Source: [Digital Curation Center](#)

Creating Data Management Plans (or DMPs)



Creating Data Management Plans (or DMPs) Definition

“Data Management Plans (DMPs) are a key element of good data management. A DMP describes the data management life cycle for the data to be collected, processed and/or generated [...].”

[H2020 Programme Guidelines on FAIR Data Management in Horizon 2020](#),
Version 3.0, 26 July 2016, p.4.

Creating Data Management Plans (or DMPs) Definition

A formal document that describes:

- the data you expect to acquire or generate during the course of a research project,
- how you will manage, describe, analyze, and store those data,
- what mechanisms you will use at the end of your project to share and preserve your data.

Creating Data Management Plans (or DMPs) Definition

- A regularly updated roadmap;
- A standardised document;
- Its content varies depending on projects' requirements and funding agencies' requests;
- Focus on data and datasets collected, created, analyzed.

Creating Data Management Plans (or DMPs) Definition

Deliverable of the project, but not a “technical” document

- It materializes the data policy of a project;
- It sums up goals and actions that will be implemented;
- It meets funder’s requirements.

Creating Data Management Plans (or DMPs)

Why make DMPs?

The main stage of the RDM

Active “management” of digital data (=> the “M” in DMP):

- An ongoing maintenance;
- An action plan in terms of data quality, technical feasibility and financial viability.

Creating Data Management Plans (or DMPs)

Why make DMPs?

The main stage of the RDM

Data management \neq Data stewardship

=

Optimizing resources for a specific purpose

- Identifying and making visible the actions to be conducted;
- Planning key stages, deadlines and critical time periods.

Creating Data Management Plans (or DMPs)

Why make DMPs?

The main stage of the RDM

Active management and digital curation

“Data curation activities enable data discovery and retrieval, maintain data quality, add value, and provide for re-use over time. This new field includes representation, archiving, authentication, management, preservation, retrieval, and use.”

[Digital Humanities Data Curation](#)

Creating Data Management Plans (or DMPs)

Why make DMPs?

The main stage of the RDM

Digital curation \neq Data storage

“Data storage is confined to simply keeping data in existence and ensuring that it can be accessed when needed. It does not necessarily entail **practices of refreshment or format migration** (essential to maintaining the data in a usable form) nor does it entail higher-level curatorial practices such as **enhancement of the data through added metadata, or migration from one representational standard to another.** Data curation thus goes far beyond the scope of data storage.”

[Digital Humanities Data Curation](#)

Creating Data Management Plans (or DMPs)

Why make DMPs?

FAIR Data

A DMP “helps Horizon 2020 beneficiaries make their research data findable, accessible, interoperable and reusable (FAIR) [...]”

[H2020 Programme Guidelines on FAIR Data Management in Horizon 2020](#), Version 3.0, 26 July 2016, p.3.

Creating Data Management Plans (or DMPs)

Why make DMPs?

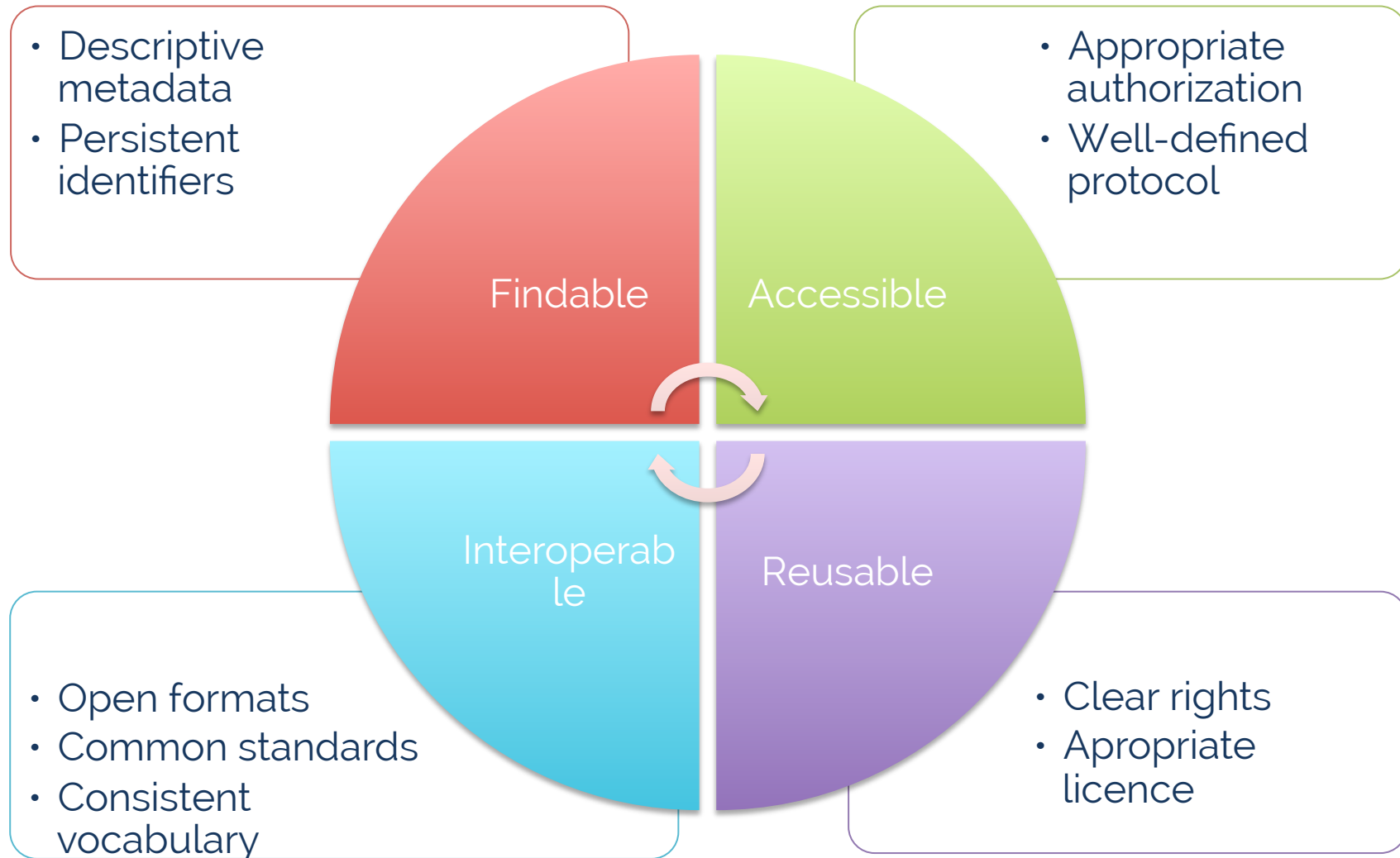
FAIR Data

- January 2014: meeting organized by the Netherlands eScience Center and the Dutch Techcentre for the Life Sciences (DTL) at the Lorentz Center in Leiden
- **FAIR principles:** “data providers and data consumers - both machine and human - could more easily discover, access, interoperate, and sensibly re-use, with proper citation, the vast quantities of information being generated by contemporary data-intensive science.” ([Force 11](#))

Creating Data Management Plans (or DMPs)

Why make DMPs?

FAIR Data



Creating Data Management Plans (or DMPs)

Why make DMPs?

To comply with funders' requirements H2020 framework

- First version of a DMP (deliverable): **first six months** of the project
- At the research proposal stage: providing a **short outline** of the data management policy.
- Since July 2016: **all the Horizon 2020 funded projects** have to provide a Data management Plan.

Creating Data Management Plans (or DMPs)

Why make DMPs?

To comply with funders' requirements
H2020 framework

Minimal requirements (initial DMP):

- A **description of data** to be generated or collected;
- The **standards and metadata** that will be used;
- The **data sharing**;
- The **archiving and preservation** (FAIR principles).

Components of a DMP

1. Information about data & data format
2. Metadata content and format
3. Policies for access, sharing and re-use
4. Long-term storage and data management
5. Budget

Crucial points to address

- Responsibility
- Results management
- Back up plan
- Intellectual property rights
- Becoming of the data after the project

Responsibility

Who does what and when?



Creates and
describes the
data



Hosting, security, ...



Data selection,
standards, mappings, ...

Information about the data // Data Collection

What data will you use?

- Reused (Cite the source)
- Created

Characterization of the data

- Raw data, derived data?
- Purpose of the data
- Volume estimation
- Type: quantitative, qualitative, survey data, experimental measurements, models, images, audiovisual data, samples, etc.



Processing of the data

Technical details on the operations that will be performed

Information about the data

// Datasets management

For each dataset, the DMP should give minimal information:

- Reference & name (Identifier for the dataset to be produced.)
- Description
 - Description of the data that will be generated or collected,
 - its origin (if collected),
 - nature & scale,
 - whether it underpins a scientific publication.
 - to whom it could be useful,
 - Information on the existence of similar data and the possibilities for integration & reuse.
- Which formats/standards are used for this data?

Datasets management: Things to avoid

File naming :

Incomprehensible names



myFile_fdslkfskj_lol(23).txt

Unknown or variable naming rules



DB-backup-20october2016.txt



database-save-20161109.txt

Special characters

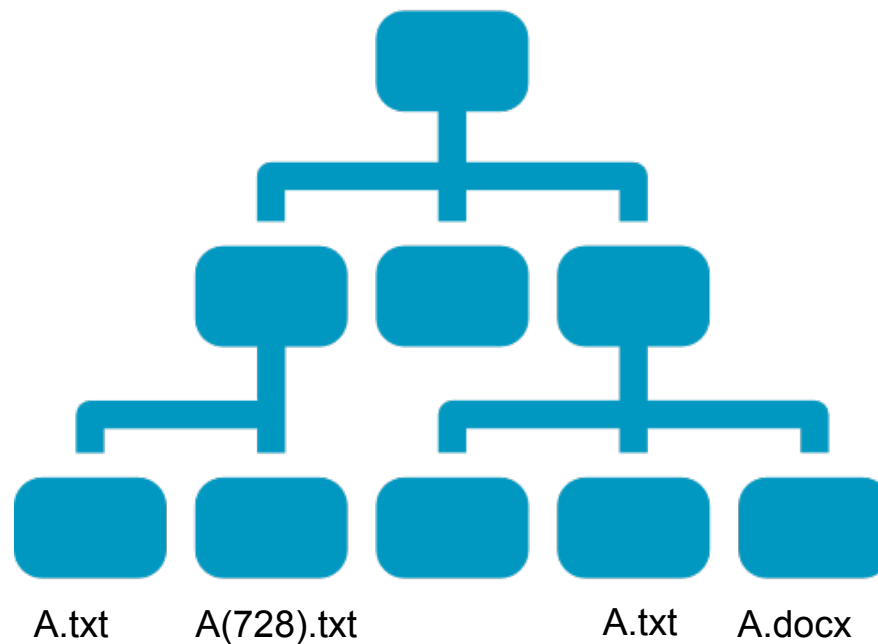


/</? .txt

Datasets management: Things to avoid

Files stored

- in several places
- in several formats and versions
- in proprietary formats



Data collection : Take-home message

- Formats
- Quality control process (How you will be sure that the created data is complete and accurate)
- Cite existing data you use
- Short term data management
- Identify Responsibilities for each task

Description and metadata

Big question : Is the data understandable by an outside researcher?

```
meaningfulData.txt
1 |504b 0304 1400 0008 0000 e37e 7a45 5ec6
2 |320c 2700 0000 2700 0000 0800 0000 6d69
3 |6d65 7479 7065 6170 706c 6963 6174 696f
4 |6e2f 766e 642e 6f61 7369 732e 6f70 656e
5 |646f 6375 6d65 6e74 2e74 6578 7450 4b03
6 |0414 0000 0800 00e3 7e7a 4531 060a 4acc
7 |b700 00cc b700 0018 0000 0054 6875 6d62
8 |6e61 696c 732f 7468 756d 626e 6169 6c2e
9 |706e 6789 504e 470d 0a1a 0a00 0000 0d49
10 |4844 5200 0000 b500 0001 0008 0200 0000
11 |7a41 a08c 0000 b793 4944 4154 789c ecbd
12 |7794 1dd7 991f 5839 bc7a 39e7 f73a 2774
13 |444e 4426 0112 cc39 48a4 3492 663d 9e99
14 |9df1 786c 8de7 d867 ffb0 f7ec eed9 b3c1
15 |bbeb 91d7 33f6 2465 8aa4 448a 3980 0449
16 |80c8 a181 6e74 a373 7c39 a78a 7b2b bd7e
17 |0049 88dd 2224 6ab6 3ff2 345e 55dd 54f7
18 |7e75 eff7 bb5f b898 2449 d03a add3 e710
19 |f6db 6ec0 3a7d a569 9d3f d6e9 56b4 ce1f
20 |eb74 2b5a e78f 75ba 15ad f3c7 3add 8ad6
21 |f963 9d6e 45eb fcb1 4eb7 a275 fe58 a75b
22 |d13a 7fac d3ad 689d 3fd6 e956 b4ce 1feb
23 |742b 5ae7 8f75 ba15 adf3 c73a dd8a d6f9
24 |639d 6e45 abe6 0f49 9260 1896 14b3 0044
25 |f901 2e6f 47cb d6e9 ab40 abe3 8f3a 7388
26 |a208 2388 7c29 df14 2108 8675 5e01 8fc0
27 |2582 c0f5 c4ca df95 343a 4b01 0693 190b
28 |5e67 b2af 30ad 8e3f ea63 094b 62a9 5435
29 |328c 2889 ead0 0a82 001e a128 8a20 0860
30 |1141 1061 30fc 08a4 3104 0489 ca7f e026
31 |8ca8 aca0 3104 482c 015e 11d5 f211 95b1
32 |a075 7ef9 6ad0 5ae6 0fbe 567a f79d 7744
33 |ca7e e7fe 5d12 579b 5b8e fb02 415c 9d4e
34 |6078 ecd2 a92a e1ee eb8c ca73 0ca4 ce1a
35 |70a9 90e3 61dc 6234 70b5 4aae 54b1 592d
36 |8bf3 7366 ab2d 9f2b 8442 4185 21e4 f261
37 |bd0a 4861 91db f1c2 ebb4 2a5a bdfc 01f2
38 |60c8 f4cc 7cf7 a6e0 fb6f bcc5 496c 3a95
39 |9210 ca6a c4ae 8cce dcfb f8d7 681a 3bf1
40 |e1b1 8951 0b81 119d 035b c84a 6c78 2a51
41 |cdce 8e4f 27fb 8736 c26c 6ee4 dab8 c31b
42 |31e3 ec95 b1f9 ddfb 0f2c cc4e b2bc 9449
43 |2658 01f2 47da 766e ed07 ab17 8420 ebdc
44 |f155 a055 f387 fc89 73ac 8461 388a 15b9
45 |5259 1039 4e20 106e 6a72 3196 8851 46b3
46 |4140 ae5d b8dc 73c7 1dd5 7c6c 6276 2148
47 |148a c542 255f 1110 71ea fa78 319b 09b4
48 |3619 103c 1e5f 2270 7cfc ea30 cff3 4613
49 |5dc8 e7e6 67e6 d22c 02f8 4399 3ed6 e78f
```

Documentation and Metadata

Metadata to describe your datasets

- Put a context around the resource :
Information about the data creation (who, when, why, how) and its use (where are they stored and what are they used for)
- Machine readable
- Standardized: DublinCore, DataCite Metadata Schema
- Automatically or manually captured
- Stored in databases, text files, or as headers in your files (Cf `teiHeader`)

+ Document your practices

- The vocabularies used to describe the data
- If you use standards or home made format for your data.
- The terms definitions and any intern conventions for description and indexation

Example : DataCite metadata standard

Datacite is a consortium of several libraries and research institutes that provide Persistent identifiers (DOIs) for research data and a metadata format to describe them.

Main fields:

- Identifier
- Creators
- Titles
- Publisher
- Publication Year
- Resource type
- Format
- Subjects
- Languages
- Version
- description



Formats

Simple equation:

Open

+ Interoperable

+ Well spread in your research community

STANDARD

Coming soon: The Parthenos Standardization Survival Kit



Standardization Survival Kit

Supporting research data modeling and
management for Arts and Humanities

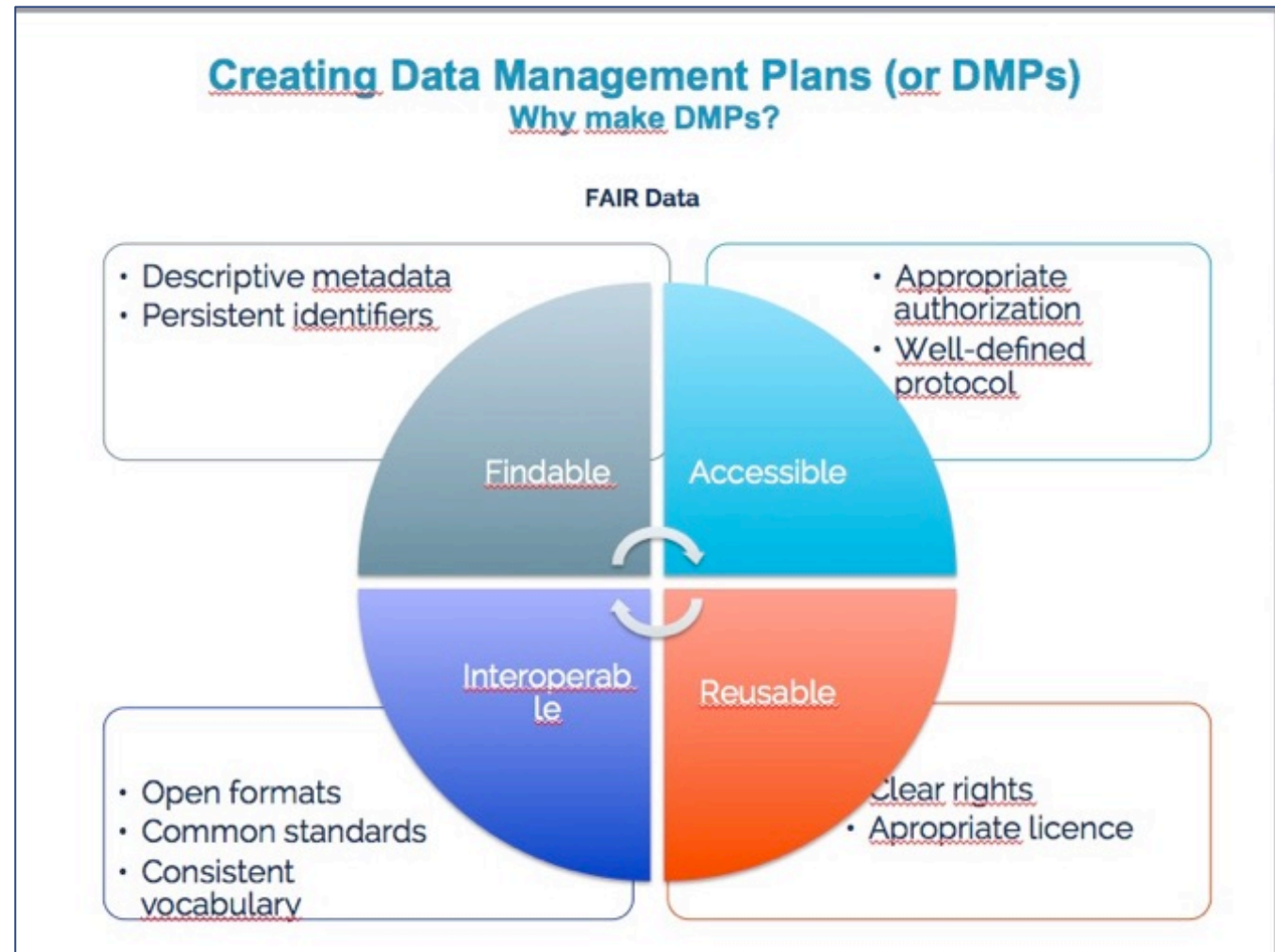
A reference environment covering
digital research scenarios in the Arts
and Humanities

It provides you with reference material about standards and their use, such as bibliographic sources, available documentation or transformations tools.

The research scenarios gathered here will serve you as examples to give you some insight on how to use standards in your own similar project.

Formats and standards: Take-home message

Remember this slide?



Storage and Backup

How will the data be stored and backed up during the research?

Anticipate incidents

Who is responsible?

How will you manage access and security?

If sensitive data (personal), adopt appropriate security measures

Questions : Where? Frequency of backups? How many copies? How many server space? Crypting? Costs? Restoring plans?

Storage = Budget + anticipation

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

Some selection criteria:

- Anticipate the futures uses and reuses
- Legal or policy aspects
- Potential value
- Consider the ratio cost/benefit

Preservation of the datasets **and also** the associated metadata, the software and algorithms used.

The European Code of Conduct for Research Integrity demands to archive primary and secondary data for a « substantial period » (European Science foundation, 2011)

should I consider storing my data in a long term archive?



are the data unique? (can be produced just once)



NO



does it involve a considerable amount of time/money to reproduce the data?

NO



any obligations to preserve your data for the long term?

NO

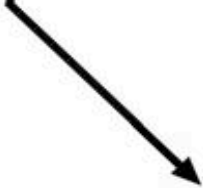


short term storage only

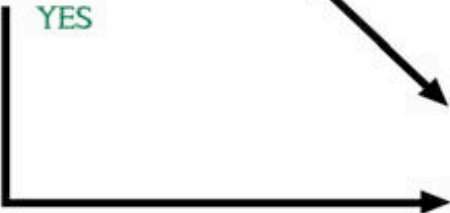
YES



YES



YES



long term archiving

Preservation (+ sharing): Data repositories



- Make data available for reuse (Harvesting, API, ...)
- Citability
- Visibility
- Transparency
- Links to papers
- Preservation

Trusted repositories

Data seal of approval

— Seals Acquired Around the World —



Storage & preservation: Take-home message

Storage → budget

Preservation → sharing

Data access and sharing

Description of how data will be shared,

- access procedures
- embargo periods (if any)
- outlines of technical mechanisms for dissemination & necessary software and other tools for enabling re-use
- definition of whether access will be widely open or restricted to specific groups.
- Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.).

If the dataset cannot be shared, give the reasons why (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related)

Ethics and intellectual property rights

Who are the right holders ? For which data?

How will you manage any ethical issues?

- Consent
- Privacy
- sensitive data

Licensing : take into account the funder policy.

Consequences on long time preservation: For example, patents data should be stored indefinitely

Sharing

- Underlying data of a scientific paper
- Data paper
- Research data repository
- Project website

“Where possible, contributors should also be uniquely identifiable, and data uniquely attributable, through identifiers which are persistent, non-proprietary, open and interoperable (e.g. through leveraging existing sustainable initiatives such as ORCID for contributor identifiers and DataCite for data identifiers).”

Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020 (https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf)

Data papers



<http://booksandjournals.brillonline.com/content/journals/24523666>

Journal of **open** archaeology data

<http://openarchaeologydata.metajnl.com/>

A **datapaper** is a scientific publication whose main goal is to describe a dataset or a group of datasets, more than analysis or research results, and to give access to the described data.

Coming soon (2) : Share your data using the...

CULTURAL HERITAGE

DATA RE-USE

CHARTER

✘ Impossible d'afficher l'image. Votre ordinateur manque peut-être de mémoire pour ouvrir l'image ou l'image est endommagée. Redémarrez l'ordinateur, puis ouvrez à nouveau le fichier. Si le x rouge est toujours affiché, vous devez peut-être supprimer l'image avant de la réinsérer.

*Offering a
comprehensive
framework*

including all
aspects relevant to
Cultural Heritage
data re-use

*Making
Cultural
Heritage Data
easier to
access, more
sustainable,
reproducible
and citable.*

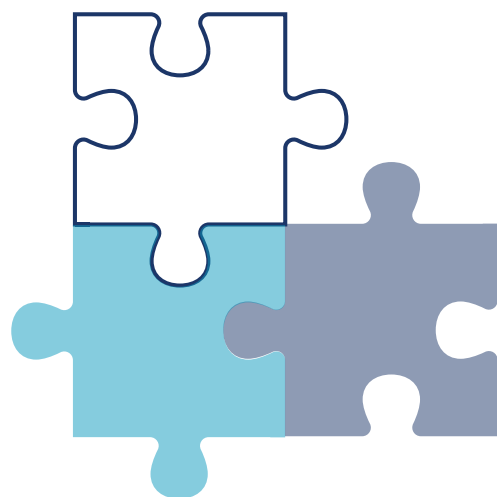
✘ Impossible d'afficher l'image. Votre ordinateur manque peut-être de mémoire pour ouvrir l'image ou l'image est endommagée. Redémarrez l'ordinateur, puis ouvrez à nouveau le fichier. Si le x rouge est toujours affiché, vous devez peut-être supprimer l'image avant de la réinsérer.

✘ Impossible d'afficher l'image. Votre ordinateur manque peut-être de mémoire pour ouvrir l'image ou l'image est endommagée. Redémarrez l'ordinateur, puis ouvrez à nouveau le fichier. Si le x rouge est toujours affiché, vous devez peut-être supprimer l'image avant de la réinsérer.

*Supporting
collaboration*

between all those
working with and
on digital data
originating from
Cultural Heritage
Institutions

Which digital data?



For which actors?

Scholars



*Research
Institutions*



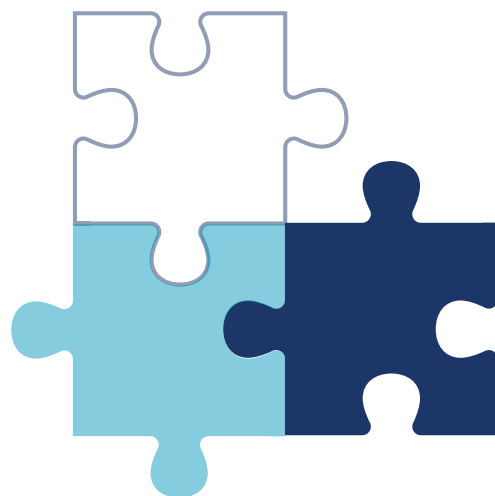
*Data Hosting
Bodies*



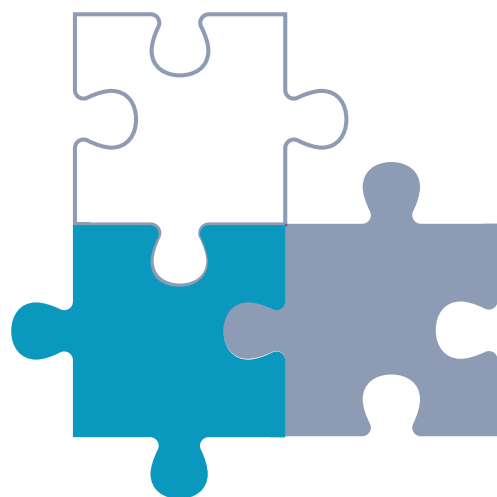
*Cultural
Heritage
Labs*



*Cultural
Heritage
Institutions*



Which benefits?



Register according to your personal or institutional profile



Get in touch with the cooperation partners and collections relevant to your activities



Gather information on relevant topics such as licensing



Gain visibility and recognition in the international research ecosystem

IPR and sharing: take-home message

Open Data Citation for Social Sciences and Humanities

The companion blog to DARIAH's Humanities at Scale Winter School in Prague: 24th-28th October 2016



To sum up

Making a DMP is defining how the data, within a project, will be

- **Described**
- **Shared**
- **Protected**
- **Preserved**

A DMP contains:

- A data lifecycle description (including long term preservation)
- A data description
- A description of the data policy
- The associated costs

A DMP helps at secure and perpetuate data.

Very strategic, but not technical

To sum up : When and why?

When?

Before the first data are created

Regularly updated

Why?

Funders wants it

Research good practice

Who?

Team work

To sum up: A research good practice

Formalize inside a unique document a set of elements and informations useful for the project monitoring and for a good management of the results.

- Understand the data
- Long-term research is easier
- No work duplication
- Underlying data is more accessible
- Research more visible : better citabilty

DMP for PHDs

PHD candidates can profitably create a DMP for the same reasons

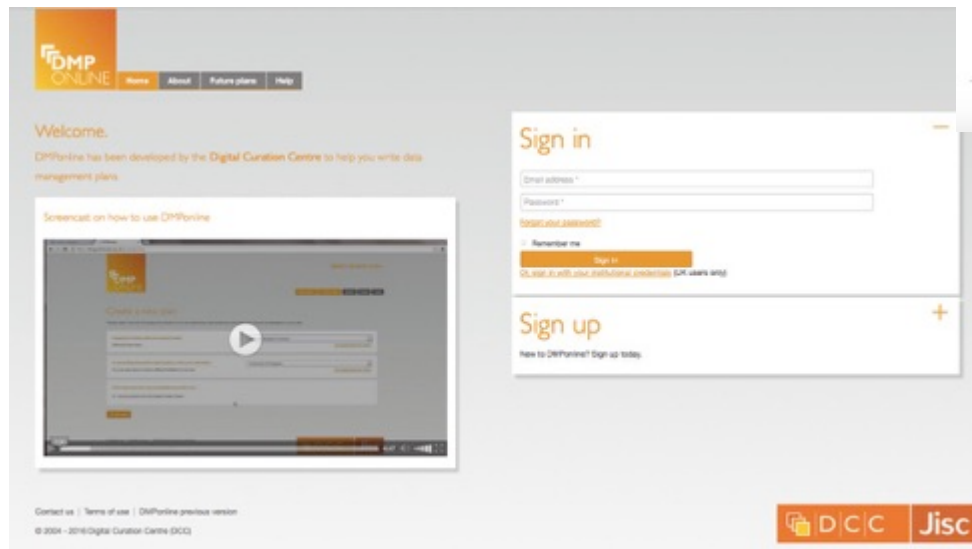
- Good practice for PHD students
- Make available key data for higher education and research
- Share data together with the thesis as underlying data

Final take-home message : The DMP Aide-mémoire

- Is there a model required by the institution/funder?
- Who will contribute to the DMP (team members, partner's projects) ?
- Who can help (documentation professionals, IT, ...)?
- Who will use the DMP?
- Use of an online tool?
- Come quickly with a first version
- Updates: required and/or desirable milestones
- Final version
- Identify datasets

Appendix : DMP tools

- Organize work in common
- Deal with the continual updating
- Choose the tool according the project specification? (e.g. confidentiality)



dmptool.org

dmponline.dcc.ac.uk

Appendix : DMPOonline (British Digital Curation Centre)

<https://dmponline.dcc.ac.uk/>

Originally based on the british needs

H2020 template, amongst others, and generic template

Online and offline

Connexion with DMPTool (US) in progress

Appendix : DMPonline Exercise

- 1) create an account
- 2) choose a model
- 3) Create and share a plan
- 4) Identify a dataset
 - Definition criteria of a dataset
 - Reasoning of the decision (reproducibility, cost, ...)
- 5) Others datasets? (granularity, strategy and concrete practice, impact)
- 6) Commenting fonction
- 7) Export

Let's try...

