Tackling Research Data Management challenges with FAIRness

Giulia Caldoni  giulia.caldoni2@unibo.it
Bianca Gualandi  bianca.gualandi4@unibo.it
Mario Marino  mario.marino6@unibo.it

Research Services Coordination Unit Research Services Division (ARIC)
University of Bologna
RDM Challenges with FAIRness

- Data need to be Findable, Accessible, Interoperable and Reusable (FAIR): need acknowledged by RDM communities
- FAIR principles can guide and be beneficial to researchers: it is not trivial to make them understand how
- RDM take place in a context of specific and restrictive ethical and legal requirements
- FAIR principles and DMP are mandatory in many funded programmes, including Horizon Europe

Researchers should be aware of how all these topics intersect the different phases of data lifecycle, and support for researchers should be tailored on this need for a global and integrated overview

EOSC Symposium, November 2022, Prague
In our experience across different research domains, researchers often face the same stumbling blocks. To support them, we have developed a RDM Decisional Tree starting from the fundamental bricks of the data lifecycle and posing a series of questions to help researchers navigate:

1) the domain specific nature and origin of the data they are handling;
2) Privacy/Ethics requirements (e.g. GDPR);
3) Intellectual Property Rights;
4) active data storage;
5) long-term deposit and preservation.

With all these questions in mind, we support researchers in finding answers by supplying a Data Lifecycle Model, including a series of possible actions and resources to protect/organize/preserve their data.
This diagram proposes a data lifecycle model inspired by the University of Virginia Library’s model (https://guides.lib.virginia.edu/c.php?g=515290&p=3522215).

Source: https://zenodo.org/record/7249051#.Y1pAi3ZBxPY
Research Data Management Decision Tree

Decisional tree to help researchers ask, and answer, questions they did not know they had. Starting from the fundamental bricks of the data lifecycle, integrates a series of questions aiming to encourage the researchers address some major attention points, encompassing issues that range from privacy/ethics requirements to IPR legislation and FAIR principles.

EOSC Symposium, November 2022, Prague
RDM Decision Tree - Planning

NEW PROJECT

DATA IDENTIFICATION

1) Qualitative/quantitative data?
2) Data format?
3) Data Size?
4) Data creator/curator?
5) Purpose of the dataset in the context of the project?

REUSE EXISTING DATA

1) Origin of the data?
2) Data licence?
3) Terms and conditions for reuse?
4) Any re-use of personal data from previous projects or activities is required?

GENERATE NEW DATA

1) Research agreements needed to ensure that data collaboratively collected are available to all of the participating researchers.
2) Are people involved in the research?
3) Which categories of people (minors, disabled, migrants, employees, etc.) are involved?
4) Which categories of personal data need to be collected and processed?
5) Complex processing operations/processing of personal data on a large scale/systematic monitoring of a publicly accessible area on a large scale are required?
6) Keep in mind the principle of Data Minimization
7) Collect informed consent for data use/sharing/preservation from research participants
8) Potential risks for participants or researchers (e.g., social stigmatisation, persecution, etc.)?
9) Possibility of incidental findings during research?
10) Development/dissemination/use of Artificial Intelligence could raise ethical concerns regarding human rights?
11) Ethical concerns about the involvement of animals, or use of substances/processes that may harm the environment, animals or plants?
12) Evaluate the possibility of Data Misuse and consider how to prevent it
RDM Decision Tree - Handling

EOSC Symposium, November 2022, Prague
RDM Decision Tree - Depositing

1) Be sure to be compliant with IPR/privacy/ethical regulations before considering deposit
2) Evaluate if embargo or restricted access to data are appropriate
3) Choose a trusted repository (PID, metadata)
4) Prepare documentation and read-me files describing the dataset
5) Choose the appropriate license
6) Metadata of the dataset should ALWAYS be public

DATA DEPOSIT

SHARING WITH OTHER RESEARCHERS

LONGTERM AND SECURE PRESERVATION BUT NOT PUBLICLY AVAILABLE

Legend:
- DATA MANAGEMENT
- INTELLECTUAL PROPERTY RIGHTS
- PRIVACY
- ETHICS
Take Home Messages:

- RDM Decision Tree is useful also as a self-standing tool to support researchers. It improves researchers' awareness on the importance and value of research data

- It helps the researcher to think about both their inbound (IPR, privacy, ethics) and outbound data, and to make upstream decisions with a view to long-term preservation.

- It highlights that FAIR RDM implies a plurality of supporting expertise and a multitude of services, resources, tools

- RDM DT is a tool that can be used in any context, it might also need to be integrated depending on the situation (e.g. funded project, different funders requirements, etc.)

- We are collecting feedback from researchers to identify further critical points and integrate the Tree

- The feedback gathered so far from researchers has shown considerable appreciation for the tool provided

EOSC Symposium, November 2022, Prague
Thank you!