universität wien

Who is afraid of the Data Management Plan?

Explorations of DMPs as useful tools in the research process.

Dr. Katja Mayer, Science and Technology Studies, University of Vienna

Image by: **bekesijoe**
https://de.toonpool.com/cartoons/data%20loss_5574

universität wien

What is a Data Management Plan?

- A structured guide (document or online tool) that covers the entire lifecycle of data
- A "living" document (to be adapted and updated during the project)
- A tool for efficient data management (not data management itself!) across all research phases and beyond
- A collection of questions on specific areas:

  Appropriate formats, standards, documentation, storage, archiving, and possibly about legal advice or technical assistance (ev. financial impact)

universität wien

Data Management Plan – Why?

- ✓ Plan, manage and reflect your data practices
- ✓ Make research verifiable, or even replicable (e.g. data papers are much easier to write with a good DMP at hand)
- ✓ Make your data understandable, findable, accessible, interoperable and reusable (yourself, research partners, research subjects and others)
  - ✓ Increase reuse potential in general
  - ✓ Increase the data benefits for research partners and research subjects
  - ✓ Prepare data for evidence based decision making
- ✓ Comply with funder mandates

DMP – What it looks like

There is a difference between

- the DMP that you put in your project proposal
  - 2-3 pages
  - A brief outlook with the most important information
  - The point here is to demonstrate preparation and thoughtfulness to funders and reviewers.
- the DMP that you use in your actual project
  - Up to 150 pages (my experience)
  - A dynamic, living document, covering most aspects of data work
  - A basis for planning and discussion
  - A tool to generate information about your work you could share with research partners and researchers (e.g. metadata of your project, basis for data papers, ...)

Elements of a DMP

Administrative information

- Project name, data originator, data owner, data controllers, contributors, contact, funding program, person IDs, etc.

Project management information

- Responsibilities for the various data practices (data controller, security, contact point …)
- Costs and resources associated with research data management

Data details

- Formats, types, metadata and standards, storage, backup and security, long-term archiving, data sharing during- and post-project use of data, legal and ethical issues

universität wien

Core Elements of a DMP

Data description	• What data and how is it generated / collected? • How will you structure the data and handle versioning? • Who is the target audience?
Documentation and Metadata	• Metadata standards • Documentation of data: FAIR; Is the data machine-readable? • Data quality control
Data availability and storage	• sharing strategy, repositories, persistent identifier, data storage and security, backup strategy, long term preservation, curation, sustainability, costs, technical barriers
Legal and ethical issues	•Data ownership? •Are there any legal barriers to making the research data fully or partially accessible? •Licensing, reuse, derivates •Ethical barriers and considerations •Sensitive data strategy

DMP – What it looks like in my research

Description of every data set

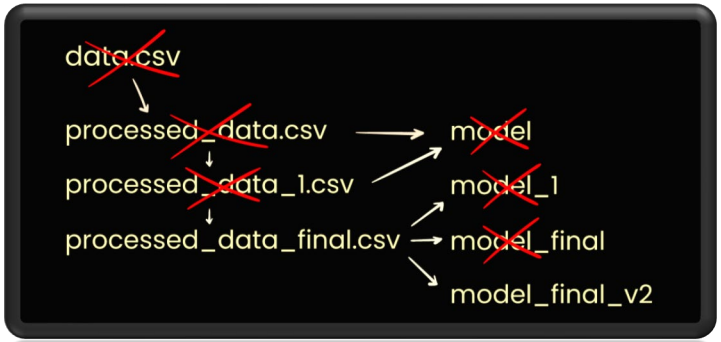
Data set reference, name, type, date, version, author
Data Set location and owner / controller
Data set description
Standards and metadata
Data Sharing locations and licenses
Archiving and preservation (including storage and backup)
Informed consent procedures and data ownership, right to be forgotten / GDPR compliance
FAIR principles
Publications

DMP Interviews with project partners

ement	Clarify storage policy	If some of the data you collect does not leave your computer/server do you plan to keep it for longer than the duration of the Co-Act project?
ement	Clarify access policy	If there is any sensitive data, how many people will have access to i and in which context?
ement	Document technical stack	What hardware or software will you use to collect data?
ement	Identify security risks	Is that hardware or software owned by you or provided by your institution?
ement	Identify security risks	If the hardware or software is provided by your institution, does it obey a security policy (admin access, firewall…)
ement	Identify security risks	If the hardware or software owned by you, do you take particular steps to protect it from malware?
ement	Identify knowledge gaps	Would you be interested in learning more about security best practices?
ement	Document data lifecycle	What file types do you store your data in? (eg, Excel, csv, something proprietary)
ement	Document data lifecycle	How do you plan to transfer data collected on the field to your main storage?
e	Identify potential data publication issues	How large do you expect your data to be? (MBs, GBs, TBs)

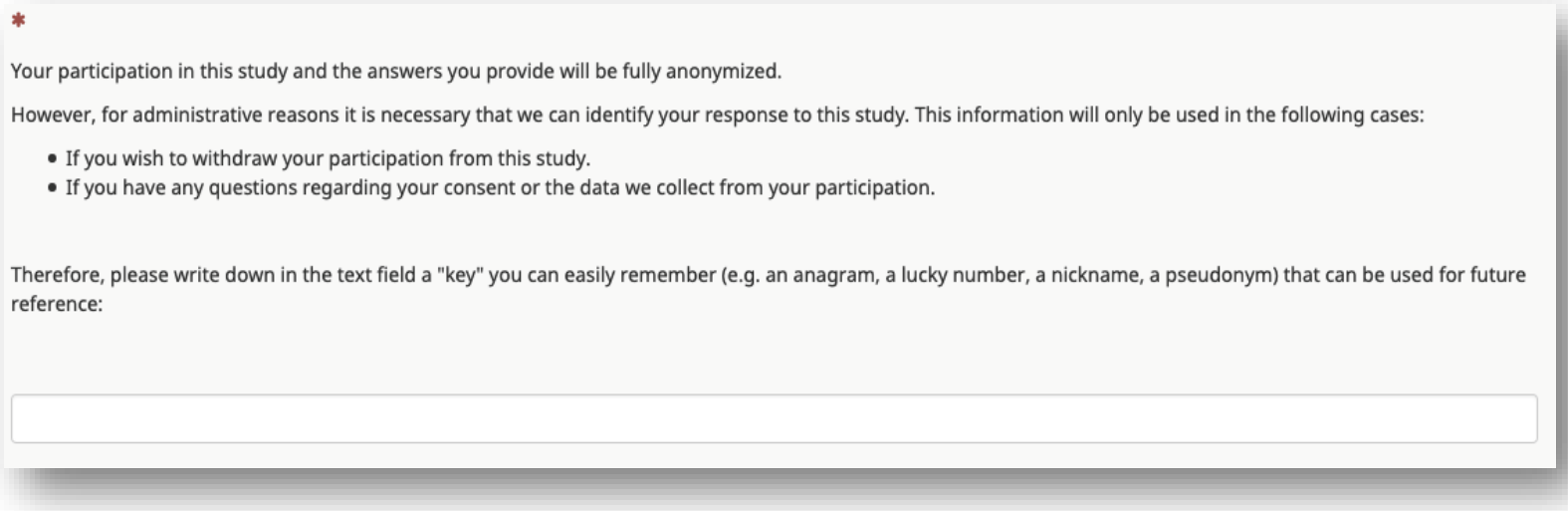
DMP and data versioning or enhancement

- Dynamic data sets, versioning
- Metadata versioning, synchronization, …
- Distribution and sharing
- Primary and secondary data use
- Data sources and their policies
- Data enhancements and derivatives
- Commercial interests and exploitation



DMP to keep track of and access for data beneficiaries

- Operationalize project sharing rules in DMP
- Connect informed consent management with DMP



*

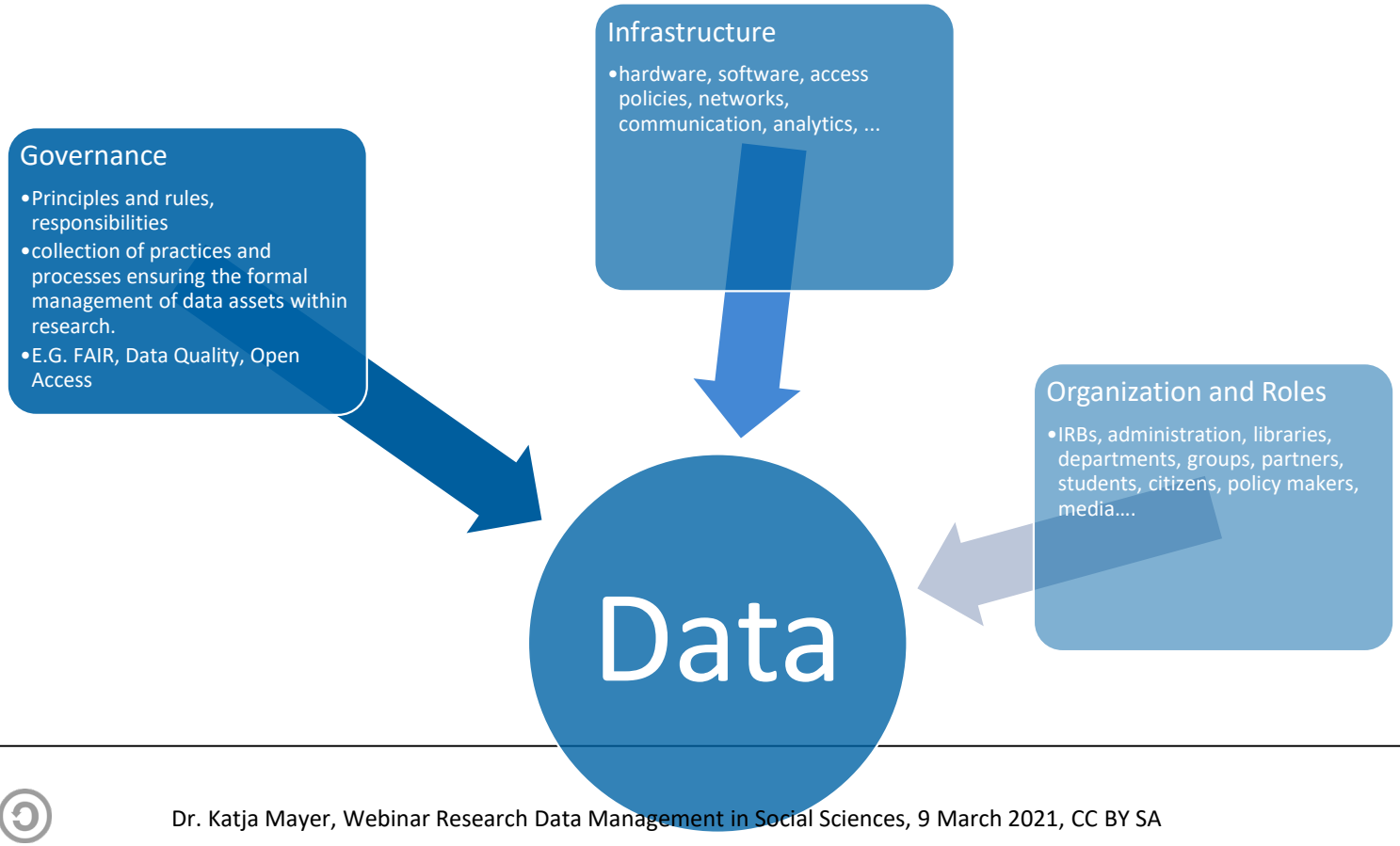
Your participation in this study and the answers you provide will be fully anonymized.

However, for administrative reasons it is necessary that we can identify your response to this study. This information will only be used in the following cases:

- If you wish to withdraw your participation from this study.
- If you have any questions regarding your consent or the data we collect from your participation.

Therefore, please write down in the text field a "key" you can easily remember (e.g. an anagram, a lucky number, a nickname, a pseudonym) that can be used for future reference:

DMP and Research Data Governance



universität wien

DMP for risk assessment

High level Scenario	Related aspects of Data Management
Poor management of data	Data Governance – roles/responsibilities
Breach of data protection laws	Information security & Privacy
Data decay of data required by compliance	Information lifecycle management
Data integration errors	Data/Information Architecture
Inability to perform regulatory / statutory reporting	Data Quality – Timeliness
Inaccurate, untimely or poor quality management	Data Quality assessment and management
Multiple data sources / lack of definitive data sources for key data	Metadata management/ Sourcing
Reliance on third parties for base data	Third party sourcing
Inadequate activity management/segregation of roles	Data Governance
lack of management review	Data Governance

https://www.dataversity.net/success-by-integrating-risk-management-principles-into-your-data-governance-policy/

universität wien

DMP to manage external data services and compliance

- Data Providers (budget, policies, …)
- Data Sources
- Transcriptions
- Translations
- Visualizations
- ….

universität wien

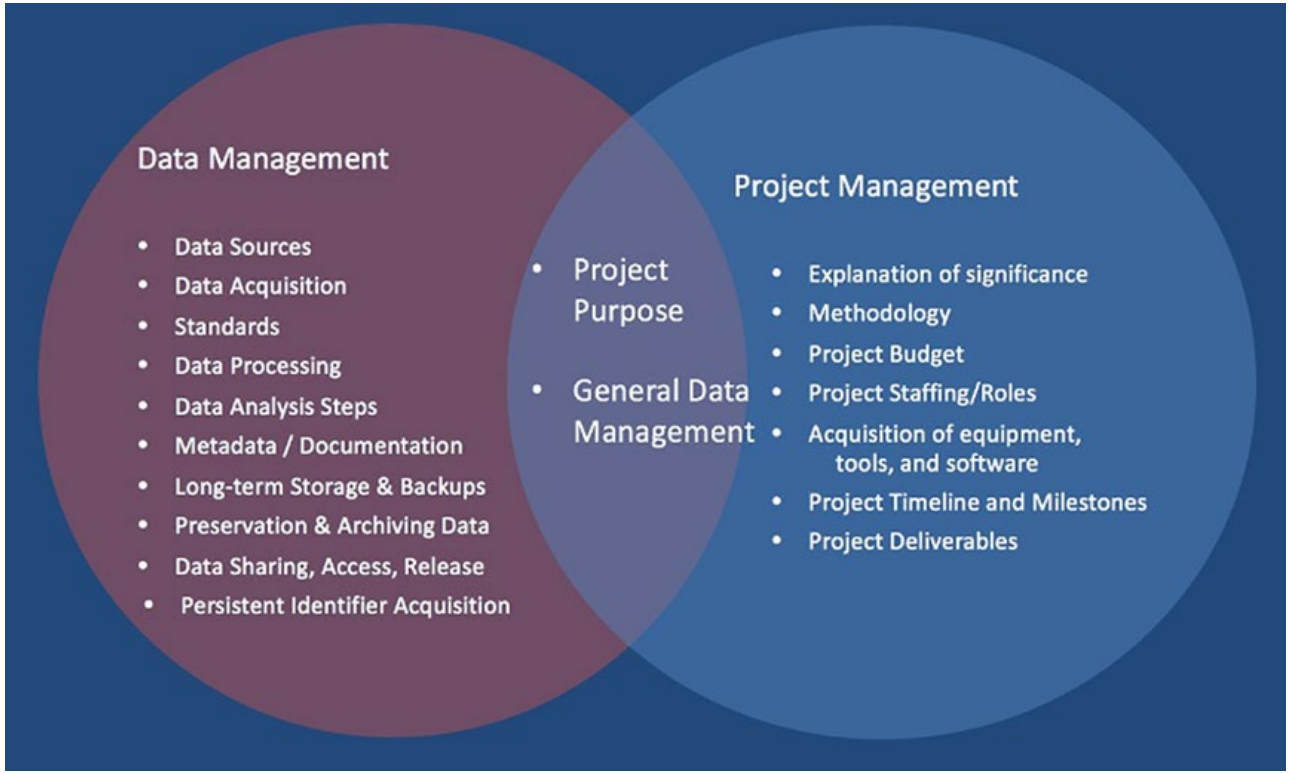
DMP to learn what can be shared

- Data too sensitive to share: what else could be shared?
  - Workflows
  - Code
  - Vignettes
  - Annotation guidelines and code books
  - Metadata
  - Annotated bibliographies
  - …..
- What data might be useful for others?
- What data to verify / replicate my research?

DMP and data papers

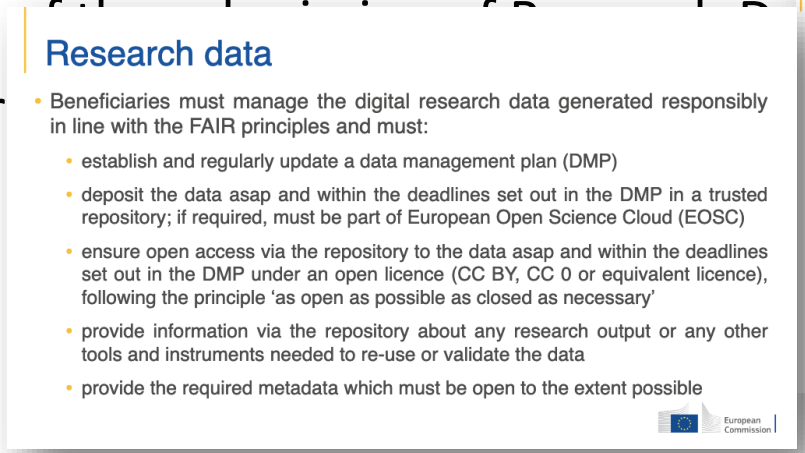
General data set properties	Data production information	Repository information and data reuse information
• Creation • Format • Type • Version • Size • Language	• Data generation and collection methods • Specimens and artifacts • Data producer • Data analysis • Missing data (and limits) • Research objectives and project information	• Repository reputation and history • Curation and digitization • Provenance • Data reuse information • Prior reuse • Terms of use / licenses

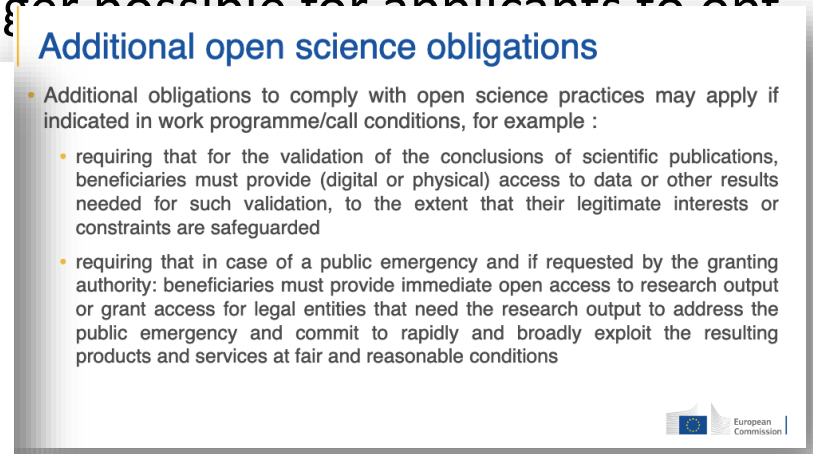
Data Management and Project Management



DMP as obligatory element in a proposal

- Austrian Science Fund FWF: https://www.fwf.ac.at/en/research-funding/open-access-policy/research-data-management
- Foundations: e.g. Volkswagenstiftung based on Science Europe template
- ERC: Finally, as from 2021 it is no longer possible for applicants to opt out
- Hor





Screenshot from slides presented on the Stakeholder workshop: Novelties in Horizon Europe

universität wien

Further information

- Science Europe https://www.scienceeurope.org/our-priorities/research-data/research-data-management/
- Cessda Training: https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide
- Data Steward Wizard: https://ds-wizard.org/
- Digital Curation Center How to Develop a Data Management and Sharing Plan and DMP Tool
- "Ethics for researchers" published by the European Commission or "The European Code of Conduct for Research Integrity".

universität wien

Thank you!

katja.mayer@univie.ac.at
@katja_mat