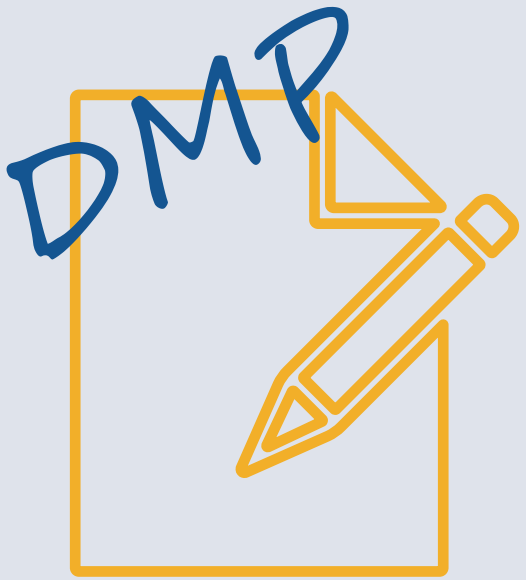
Data Management Plan

Irena Njezic

Data Management Plan



What is a Data management plan?

Why is it important for one’s research?

What are elements of a DMP?

DMP tools

What is a DMP?

A Data Management Plan is a document specifying how research data will be handled both during and after a research project.

DMP is a living document

Mandatory

Why is DMP important for one’s research?

Keeps the data safe

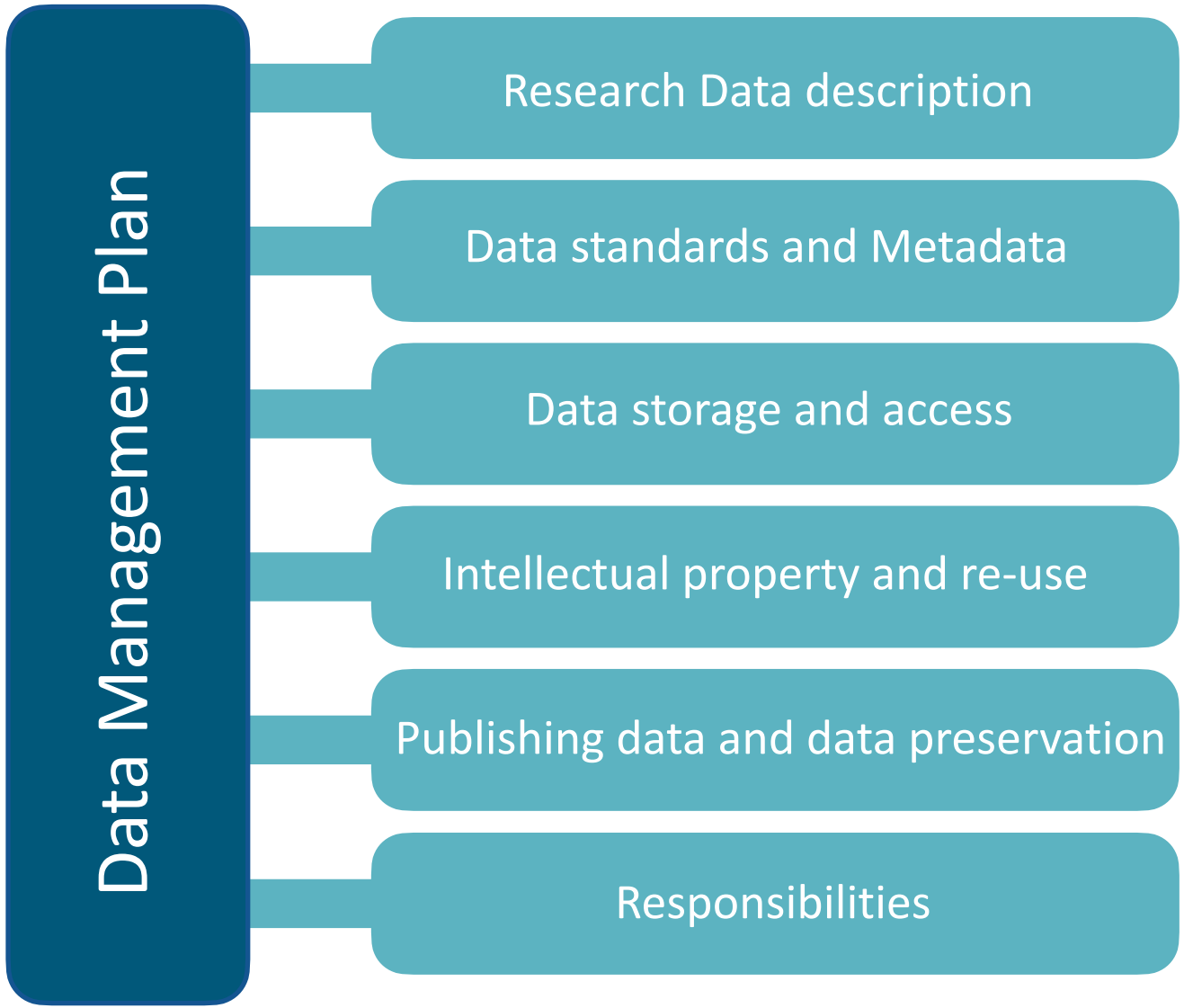
Helps researchers organise their data

Helps avoid wrongful accusations (Climategate)

WIREs Clim Change 2012, 3:289–295. doi: 10.1002/wcc.168

It encourages the reuse of data

Funders require researchers to write a DMP



What are elements of a DMP?

DMP tools

Templates

https://www.cms.hu-berlin.de/de/dl/dataman/muster-dmp-h2020-v3

https://www.cms.hu-berlin.de/de/dl/dataman/muster-dmp-dfg

https://www.cms.hu-berlin.de/de/dl/dataman/muster-dmp-bmbf

https://www.cms.hu-berlin.de/de/dl/dataman/muster-dmp-vwstiftung-pdf

http://www.icpsr.umich.edu/icpsrweb/content/datamanagement/dmp/framework.html

Checklists

https://www.forschungsdaten-bildung.de/get_files.php?action=get_file&file=fdbinfo_2.pdf

www.dcc.ac.uk/sites/default/files/documents/resource/DMP/DMP_Checklist_2013.pdf

http://opus.bath.ac.uk/36009/4/DMP_Guidance_for_PGRs_v1.3.pdf

http://www.forschungsdaten.org/images/b/b0/Leitfaden_Data-Management-WissGrid.pdf

Tools

DMPTool (https://dmptool.org/)

RDMO (http://rdmorganiser.github.io)

DMPonline (https://dmponline.dcc.ac.uk/)

Data Stewardship Wizard (https://ds-wizard.org/)

OpenDMP/Argos (https://argos.openaire.eu)

Source: Biernacka, Katarzyna, Bierwirth, Maik, Buchholz, Petra, Dolzycka, Dominika, Helbig, Kerstin, Neumann, Janna, Odebrecht, Carolin, Wiljes, Cord, & Wuttke, Ulrike. (2020). Train-the-Trainer Concept on Research Data Management (3.0) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.4071471

DMP Template

Data Description (Recommended)

Provide a brief description of the information to be gathered -- the nature, scope, and scale of the data that will be generated or collected.

Why this is important
A good description of the data to be collected will help reviewers understand the characteristics of the data, their relationship to existing data, and any disclosure risks that may apply.

+ Examples

Example 1:
This project will produce public-use nationally representative survey data for the United States covering Americans' social backgrounds, enduring political predispositions, social and political values, perceptions and evaluations of groups and candidates, opinions on questions of public policy, and participation in political life.

Example 2:
This project will generate data designed to study the prevalence and correlates of DSM III-R psychiatric disorders and patterns and correlates of service utilization for these disorders in a nationally representative sample of over 8000 respondents. The sensitive nature of these data will require that the data be released through a restricted use contract.

Access and Sharing (Recommended)

Indicate how you intend to archive and share your data and why you have chosen that particular option. Possible mechanisms for archiving and sharing include:

- Domain repository like ICPSR (social science)
- Self-dissemination through a dedicated website that the research team will create and maintain. *If this option is chosen, it is recommended that the data producer arrange for eventual archiving of the data after the self-dissemination period terminates and specify the schedule for data sharing in the grant application.*
- Preservation with delayed dissemination. Under such an agreement the data producer makes an arrangement with a public data repository for archival preservation of the data with dissemination to occur at a later date, usually within a year.
- Institutional repositories. Institutional repositories at academic institutions have the goal of preserving and making available some portion of the academic work of their students, faculty, and staff. Note that not all IRs have the capacity to accept and curate data.

Why this is important
Sharing data helps to advance science and to maximize the research investment. A recent paper reported that when data are shared through an archive, research productivity increases and many times the number of publications result as opposed to when data are not shared.

Protecting research participants and guarding against disclosure of identities are essential norms in scientific research. Data producers should take efforts to provide effective informed consent statements to respondents, to deidentify data before deposit when necessary, and to communicate to the archive any additional concerns about confidentiality. (See Ethics and Privacy below.)

Metadata (Recommended)

What types of metadata will you produce to support the data? Will a metadata standard be used?

Why this is important
Good descriptive metadata are essential to effective data use. Metadata are often the only form of communication between the secondary analyst and the data producer, so they must be comprehensive and provide all of the needed information for accurate analysis.

Structured or tagged metadata, like the XML format of the Data Documentation Initiative (DDI) standard, are optimal because the XML offers flexibility in display and is also preservation-ready and machine-actionable.

+ Examples

Intellectual Property Rights (Recommended)

Who will hold intellectual property rights for the data and other information created by the project?

Will these rights be transferred to another organization for data distribution and archiving? Will any copyrighted material (e.g., instruments or scales) be used? If so, how will the project obtain permission to use the materials and disseminate them?

Why this is important
In order to disseminate data, archives need a clear statement from the data producer of who owns the data. The principal investigator's university is usually considered to be the holder of the intellectual property rights for data the PI generates. Many archives do not ask for a transfer of rights but instead just request permission to preserve and distribute the data. Copyright may also come into play if copyrighted instruments are used to collect data. In these cases, data producers should initiate discussions with archives in advance of data deposit.

+ Examples

Ethics and Privacy (Recommended)

If applicable, how will you handle informed consent with respect to communicating to respondents that the information they provide will remain confidential when data are shared or made available for secondary analysis?

Why this is important
Protection of human subjects is a fundamental tenet of research and an important ethical obligation for everyone involved in research projects. Disclosure of identities when privacy has been promised could result in lower participation rates and a negative impact on science.

+ Examples

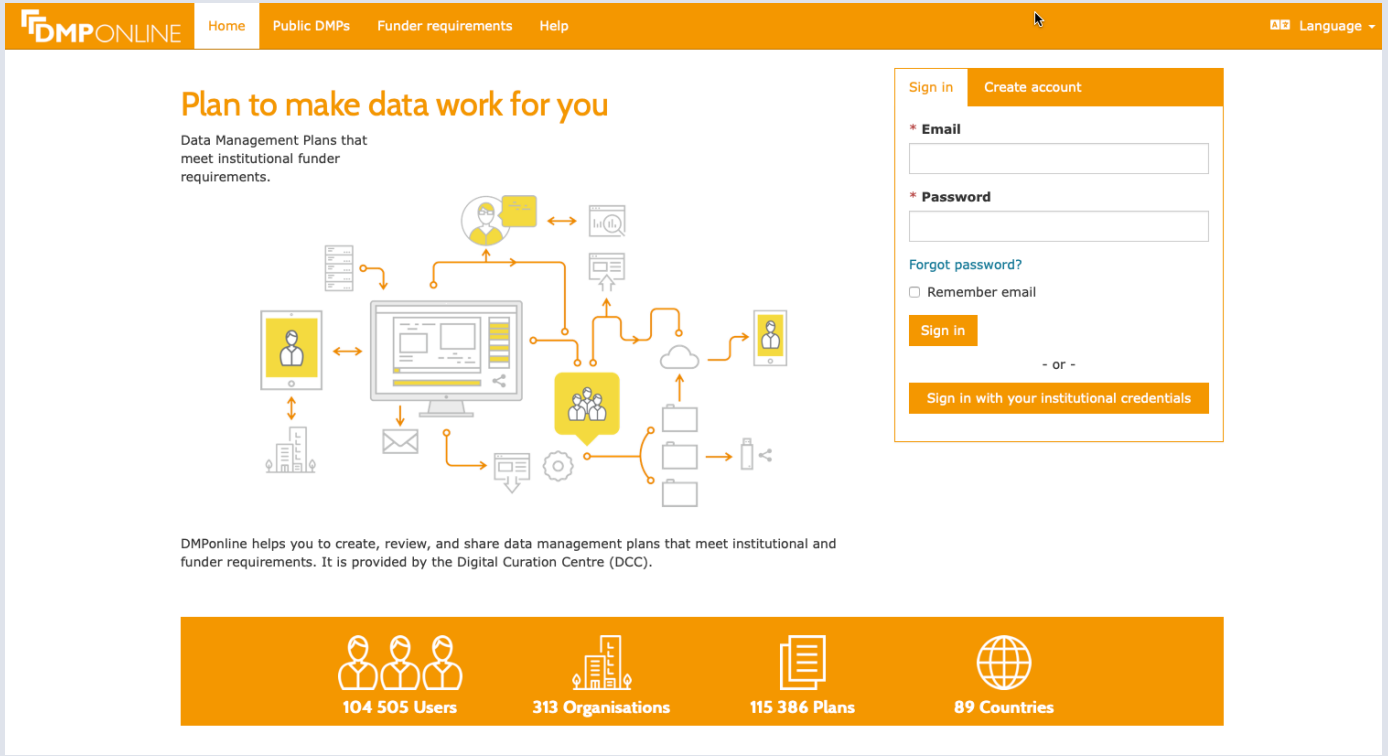
Format (Recommended)

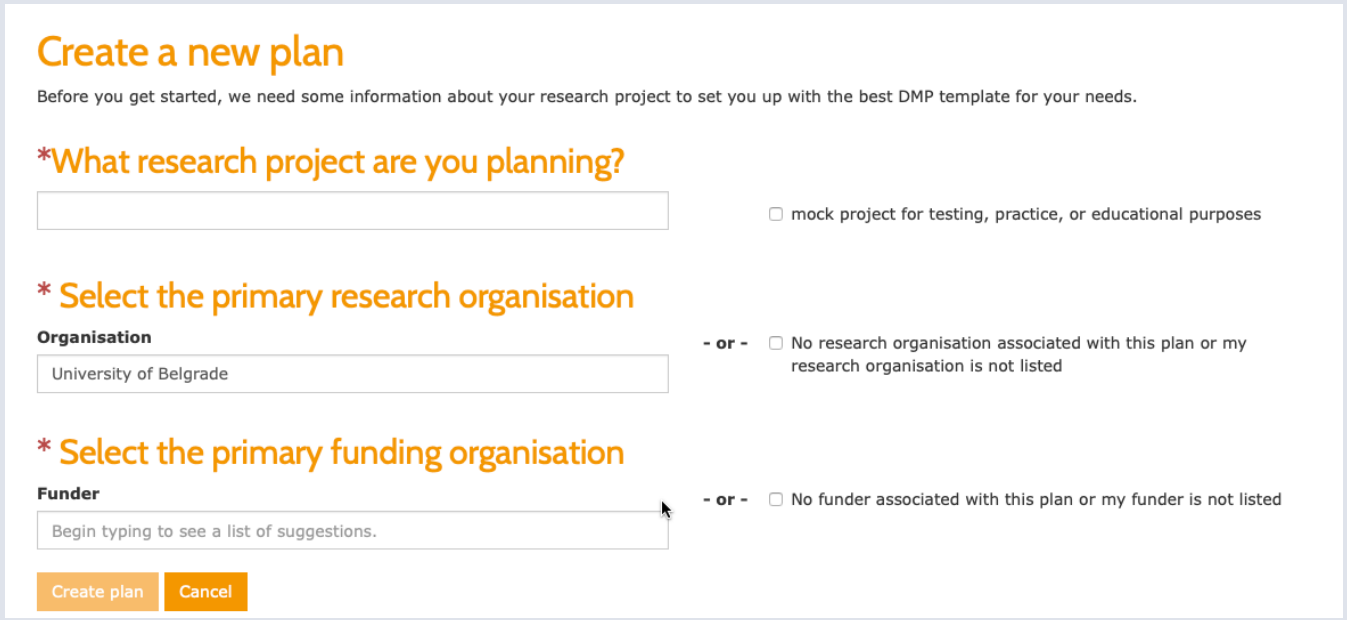
Specify the anticipated submission, distribution, and preservation formats for the data and related files (note that these formats may be the same).

Why this is important
Depositing data and documentation in formats preferred for archiving can make the processing and release of data faster and more efficient.

DMP Online

DCC - Digital curation centre
Free to use it once you create an account
Templates





https://dmponline.dcc.ac.uk

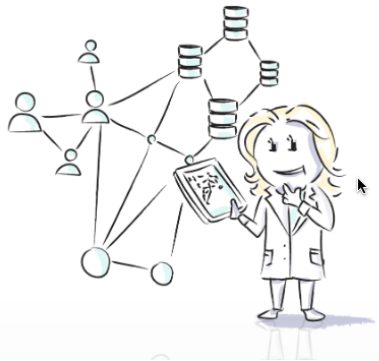
argos

Plan and follow your data

Create machine actionable DMPs.
Configure to best fit your discipline.
Link to EOSC components out of the box.
Share easily in your repository.

Bring your Data Management Plans closer to where data are generated, analysed and stored.

Start your DMP



Argos

OpenAIRE
Public DMPs
Templates
Export
Free for researchers

Sign in to account

thank you!