Drafting a data management plan for an international research project

François Vincent et Martine Peters
Université du Québec en Outaouais

CRSH SSHRC
Conseil de recherches en sciences humaines
Social Sciences and Humanities Research Council

Partnership on University Plagiarism Prevention
PUPP

What is the Partnership on University Plagiarism Prevention (PUPP)?

- Funded by the *Social Sciences and Humanities Research Council* (SSHRC) – Canadian Government
- Partnership Grant (2021-2028)
- 10 countries (Canada, United-States and Europe)
- 35 Partners (universities and associations like ENAI)
- 63 team members – researchers, collaborators (librarians, administrators, coordinators, advisors, etc.)
- A compulsory requirement by the SSHRC: a data management plan

Objective of this presentation

Presenting the steps for creating a good data management plan for an international project, as well as the issues and reflections that emerged at each of these steps.

What is a DMP ?

- A data management plan (DMP) is a document that aims to describe how you will treat your data during a project and what will happen to them after the research is completed (Alcina Guandalini, Furnival & Simionato Arakaki, 2019 ; Michener, 2015).
- Usually structured according to the research funder (Michener, 2015)
- An evolving document (SSHRC, 2023)
- Planification of the storing, using, maintaining and making available the data (Hudson-Vitale & Moulaison-Sandy, 2019)

Why have a data management plan (DMP) ?

Miksa, Simms, Mietchen & Jones, 2019

Why have a data management plan (DMP) ?

- Embrace good practices in the scientific community (Alcina Guandalini et al, 2019)
- A significant change in researcher practices and expectations of support (Krahe et al., 2020)

How did we elaborate our DMP?

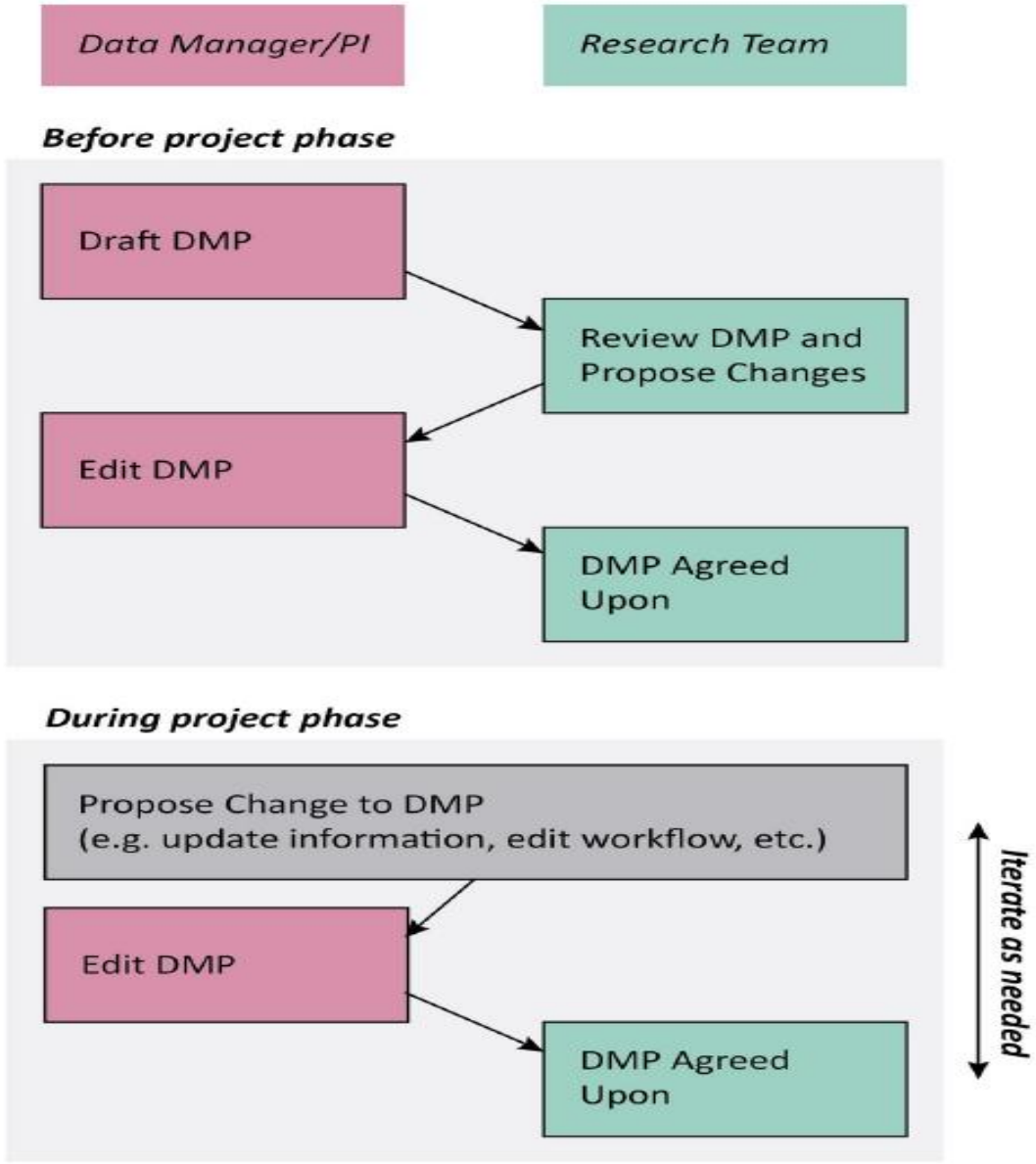
Adopted a template from DMPAssistant (most models are similar)

https://alliancecan.ca/en/services/research-data-management/dmp-assistant

8 steps

1. Data management Policy
2. Data collection
3. Documentation and metadata
4. Storage, access and backup
5. Preservation
6. Sharing and reuse
7. Responsabilities and ressources
8. Ethical and legal conformity

Process and roles for developing the DMP



Bryney, Goben and Jones, 2023

How did we elaborate our DMP?

Took almost a year of regular meetings

Impossible to start data collection before the DMP was completed

Who participated – from the beginning – PUPP director and two co-directors and scientific committee (representation of various countries)

Once first draft completed – all team members

Once second draft completed – lawyers

Final version – deans for signatures

How did we elaborate our DMP?

1. Data management policy

SSHRC — compulsory, minimalistic (go see this or that!), Canadian policy even if international project

How did we elaborate our DMP?

2. Data collection

- All the DMP is build on this section.
- Most elaborated section with all the details about instruments, methods, software, ethics, security and confidentiality, data transfer, files format, etc…
- For PUPP – adoption of the management of data, all guidelines and methods were determined, put into annexes (n=13) – coming to an understanding took forever!!!

How did we elaborate our DMP?

3. Documentation and metadata

- All information needed to use data and to eventually publish it
- To permit reproduction of the research project
- If need to give data for publication
- To share data after the end of project (Dataverse in Canada)

Examples of metadata: type of software used (Limesurvey), scales used with questionnaires, date of data collection, etc…

How did we elaborate our DMP?

4. Storage, access and backup

    Active phase of the project: where, when, how, who – will use and access the data

Biggest headache because of various countries and different ways of storing, and ethic committees requirements !

PUPP produced a step-by-step guide for procedures on how to save and store raw and treated data.

How did we elaborate our DMP?

5. Preservation

    After the project, what will happen to the data, storage and type of file format

Had to take into account the evolution of file formats over time so that the data would be accessible in the long term. Document created to transfer data from one type of file to another

How did we elaborate our DMP?

6. Sharing and reuse

    After the project, need to determine which data (not raw but anonymised) will be used by anyone… Dataverse

    Obligation in Canada

PUPP – not understood by team members outside of Canada

data but also educational resources will be shared

How did we elaborate our DMP?

7. Responsibilities and resources

    Choice of who will be responsible for the data and which committee

PUPP DMP planned– 3 committees: scientific, quantitative and qualitative and each university is responsible for their own data

In reality: data collection is centralised, anonymised and then distributed to all team members

How did we elaborate our DMP?

8. Ethical and legal conformity

   Ethic committees' (EC) approval necessary for all ethics, legal and intellectual property aspects.

PUPP – EC extremely different from one country to the next, from very complicated to none / ethics grouping or individual

- confidential forms for all team members and their students
- consent forms, minor changes between institutions

Now – application of the DMP…

The reality is very different than what was planned for many reasons:

- lots of training before the data collection started to insure uniformity
- big team means lots of people involved in data collection
- multidisciplinary team
- rigour is different from one country to the next
- language issues: misinterpretation of the instructions, translation

So the DMP is, for practical reasons (the cumbersomeness of the process), is constantly in evolution.

Conclusion

A rigorous way of designing, conducting and adjusting research

No plan is perfect, blind spots are inevitable

Doing research with a DMP is taking another step closer to integrity

Need DMP models to circulate

Références

Alcina Guandalini, C., Furnival, A. C. M., & Simionato Arakaki, A. C. (2019). Good Scientific Practice in the Drawing up of Data Management Plans. *Revista Digital de Biblioteconomia e Ciencia da Informacao / Digital Journal of Library and Information Science*, *17*, 1-19 **doi**: 10.20396/rdbci.v017i0.8655895

Briney, K. A., Goben, A. et Jones, K. (2022). Data Management Planning for an Eight-Institution, Multi-Year Research Project. International Journal of Digital Curation, 16(1), 1-9. https://doi.org/10.2218/ijdc.v16i1.799

Cheung, M., Cooper, A., Dearborn, D., Hill, E., Johnson, E., Mitchell, M. et Thompson, K. (2022). Practices Before Policy: Research Data Management Behaviours in Canada. Theory and Research, 17(1), 1-80. https://doi.org/https://doi.org/10.21083/partnership.v1i2.6779

Gibson, J. (2023). Sharing, curation and metadata as essential components of the data management plan. Information Services & Use, 43, 397-401. https://doi.org/10.3233/ISU-230220

Hudson-Vitale, C., & Moulaison-Sandy, H. (2019). Data Management Plans: A Review. *DESIDOC Journal of Library & Information Technology*, *39*(6), 322-328. DOI : 10.14429/djlit.39.6.15086

Michener, W. (2015). Ten Simple Rules for Creating a Good Data Management Plan. PLoS Comput Biol, 11(10). https://doi.org/10.1371/journal.pcbi.1004525

Miksa, T., Stephanie, S., Metchen, D. et Jones, S. (2019). Ten principles for machine-actionable data management plans. PLoS Comput Biol, 15(3), 1-15. https://doi.org/https://doi.org/10.1371/journal.pcbi.1006750

Smale, N., Denyer, G., Unsworth, K., Magatova, E., & Barr, D. (2020). A review of the history, advocacy and efficacy of data management plans. *International Journal of Digital Curation*, 15(1), 1-30. http://dx.doi.org/10.2218/ijdc.v15i1.525