

Mapping of DMP stakeholders and user needs in the project 'Data Management Plans: Support package for Norwegian higher education libraries'

Jenny Ostrop ¹, Live Håndlykken Kvale ², Svein Høier ³, Ingrid Heggland ³,
Leif Longva ⁴, and Korbinian Bösl ^{1,5}

¹University of Bergen

²University of Oslo

³Norwegian University of Science and Technology

⁴UiT The Arctic University of Norway

⁵ELIXIR Norway

The project 'Data Management Plans: Support package for Norwegian higher education libraries' aims to create a comprehensive Data Management Plan (DMP) support package with resources and guidance both for users and support personnel at Norwegian research institutions. Thereby the project follows up on recommendations on establishing a common understanding of DMP requirements and aligning DMP guidance between institutions from a collaborative DMP project by the University libraries in Bergen, Oslo, Trondheim and Tromsø published in December 2022 [1].

To establish a common ground for defining DMP outlines and guidance, the project group has conducted a mapping of DMP stakeholders in Norway. This mapping is both inspired by taking a multi-stakeholder perspective on data management plans as proposed by Kvale and Pharo [2] and the group's extensive experience in providing data management support in institutions and research infrastructures to both researchers and policy makers.

Furthermore, the project group has categorised aspects that a DMP may have to cover as different researchers and research projects can have opposing needs. In order to provide relevant guidance, the whole spectrum should be considered.

Writing a DMP contributes to making data and other outputs of research projects FAIR. Eventually, if certain prerequisites are met, the DMP itself can be considered a research object and be integrated in the Open Science ecosystem.

Defining these premises was valuable for ongoing the project work. We invite the research data community to provide feedback.

1 DMP stakeholders

A data management plan (DMP) is a structured document containing information and thoughts regarding the handling of research data throughout the entire lifecycle of a research project, with the intention of making data as FAIR (Findable, Accessible, Interoperable, Reusable) and as open as possible. A data management plan is a living document to be updated as the project develops. Researchers meet requirements from institutions and research funders to create DMPs for their projects, in order to maximise project impact and to ensure that legal and ethical aspects are considered. For researchers, writing a DMP should be an awareness-creating process and the DMP can function as a hub that connects data-related aspects in a project. Relevant guidance is essential in order to make writing a DMP a useful exercise instead of being perceived as an administrative burden.

DMP stakeholders

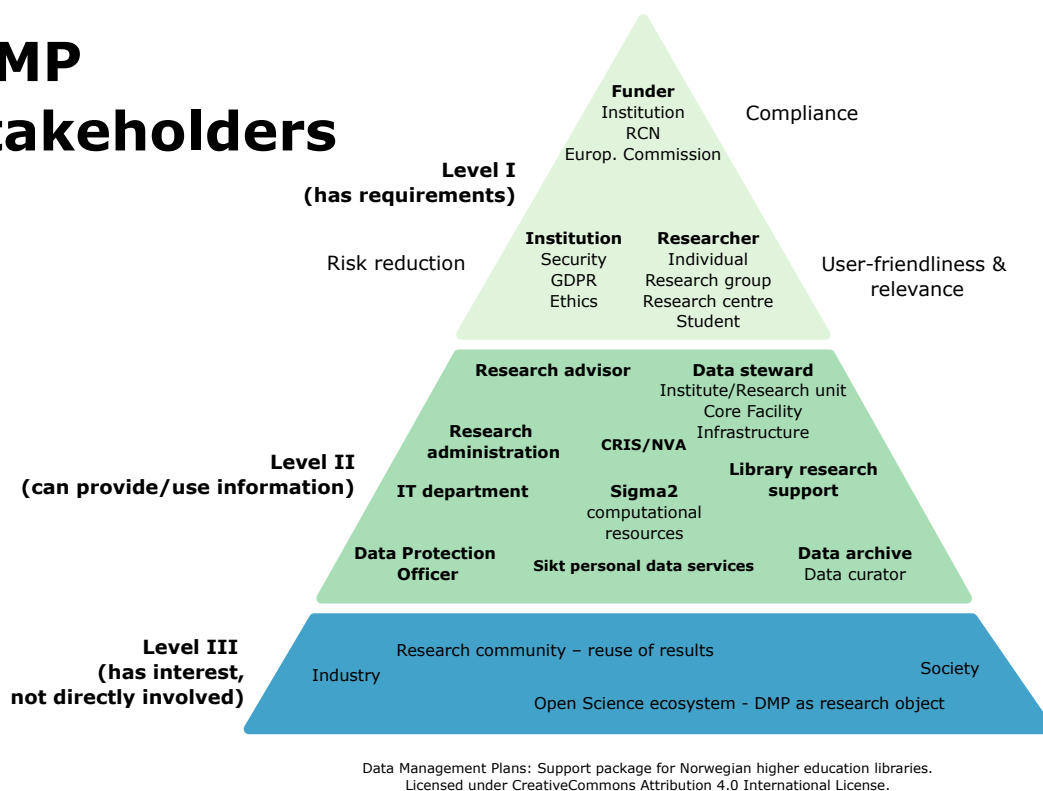


Figure 1: Stakeholder pyramid: DMP stakeholders in Norway.

Importantly, the information in a data management plan can also be valuable for third-party actors at different stages of a research project. Integrating information in DMPs with other services through machine-actionable data management plans (maDMP) is increasingly gaining attention [3, 4]. Of note, not all of the suggested interactions are currently covered by the RDA Common Standard for machine-actionable DMPs [5]. The current standard is best suited for documenting administrative information and produced datasets, and thus exchanging DMP information connected to the beginning or end of a research project. Aspects related to conducting research in the active phase, such as routines for data management and processing or legal and ethical considerations, are currently not sufficiently covered. These aspects are more difficult to standardise and closely linked to the research domain, yet potentially of interest for stakeholders and crucial for the awareness-creating function of a DMP.

In this project, we have categorised DMP stakeholders in 3 levels as illustrated in Fig. 1 as stakeholder pyramid: stakeholders that have requirements on DMP format or content (level I), stakeholders that could interact with information in the DMP (level II), and stakeholders that are not directly involved but have an interest in DMPs facilitating FAIR research data or in the DMP itself (level III).

Level I: Stakeholders who set requirements

- Researchers (employed in various structures, at different career levels): Requirements for user-friendliness and academic relevance
- Institutions: Requirements for complying with information security, legislation and research ethics
- Research funders, Institutions that require DMP: Requirements for Open Science and DMP

Level II: Stakeholders who can offer/use information in DMP

The level II stakeholders might have interests in a DMP at different times during a research project, the relevant project phases are indicated in square brackets: planning (pre/post-award),

implementation (active phase of the research: data creation/collection, data processing, preparing data for archiving), reporting (to institution/funder, e.g. using research information system).

Within the institution

- IT department: Data storage services, information security guidelines [planning, implementation]
- Data Protection Officer: Projects with personal data/special category personal data, research ethics guidelines [planning]
- Research Advisor/Research Administration: Project funding, project results, for evaluation and statistical purposes [planning, reporting]

Support personnel

- Data Steward (affiliated with research unit, core facility, or research infrastructure): Planning, data documentation, data storage, data analysis, data sharing [planning, implementation]
- Research support at the university library: Planning, data documentation, data sharing [planning, implementation]
- Data Curator (at data archive): Data sharing [implementation, reporting]

National/regional service providers

- Sigma2 (national service provider of computational infrastructure): Allocating data storage services, high-performance computing [planning]
- REK (regional ethical committees): Projects with health data [planning]
- SIKT (Norwegian Agency for Shared Services in Education and Research) privacy services: Projects with personal data [planning]
- CRISStin/NVA (research information system): Project results [reporting]

Level III: Stakeholders without immediate involvement

- Research community: Reuse of research data, transparency in the research process, overview of ongoing projects
- Open Science ecosystem: DMP as research object
- Industry: Re-use of research data
- Society: Transparent and cost-effective research

2 DMP needs

Depending on the field of research and the needs of the project, the DMP may take different forms. A DMP for a straightforward project with few participants and limited data volume will look different from a DMP for a complex international consortium project that will produce terabytes of data. A DMP for a project that analyses literature sources will look different from a project that creates many different types of data from patient samples. The project aims to address these - partially opposing - ‘DMP cornerstones’, as illustrated in Fig. 2.

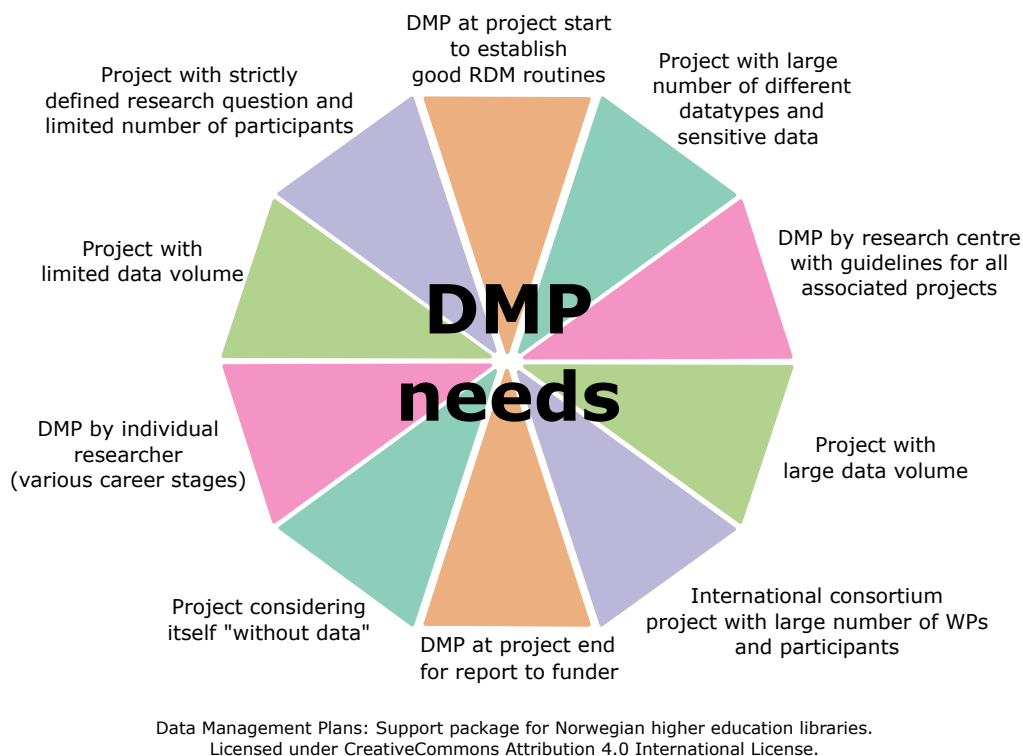


Figure 2: 'DMP cornerstones' illustrating the different needs a DMP may need to meet.

3 DMPs in relation to FAIR digital objects

Level III stakeholders have an interest in DMPs facilitating FAIR research data or in the DMP itself without directly being involved. This can both include a human reading a DMP and machines collecting DMP information. In order to be integrated in the Open Science ecosystem, a DMP should be equipped with a persistent identifier (PID) and published, preferably in a research data archive [3, 6]. In addition, standards need to be established facilitating exchange of metadata from the DMP to open science infrastructures [7, 8]. Seeing DMP in the context of linked open data [9, 10] is another step towards DMPs as integrated research objects in the Open Science ecosystem. Standards and applications are still under development and it is expected that the field will continue to evolve rapidly during the project period. Yet, our DMP guidance should encourage researchers to assign DMPs a PID and provide sufficient metadata to enable reuse by humans and machines.

References

- [1] A. Gabrielsen et al. *UB-BOTT-samarbeid om datah ndteringsplaner: kartlegging og anbefalinger*. Zenodo, Dec. 12, 2022. DOI: 10.5281/zenodo.7428542.
- [2] L. Kvale and N. Pharo. "Understanding the Data Management Plan as a Boundary Object through a Multi-stakeholder Perspective". In: *International Journal of Digital Curation* 16.1 (1 July 4, 2021), pp. 16–16. ISSN: 1746-8256. DOI: 10.2218/ijdc.v16i1.746.
- [3] T. Miksa et al. "Ten Principles for Machine-Actionable Data Management Plans". In: *PLOS Computational Biology* 15.3 (Mar. 28, 2019), e1006750. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1006750.
- [4] T. Miksa, S. Oblasser, and A. Rauber. "Automating Research Data Management Using Machine-Actionable Data Management Plans". In: *ACM Trans. Manage. Inf. Syst.* 13.2 (June 30, 2022), pp. 1–22. ISSN: 2158-656X, 2158-6578. DOI: 10.1145/3490396.
- [5] T. Miksa, P. Walk, and P. Neish. "RDA DMP Common Standard for Machine-actionable Data Management Plans". In: (Sept. 16, 2020). DOI: 10.15497/rda00039.

- [6] S. Simms et al. “Machine-Actionable Data Management Plans (maDMPs)”. In: *Research Ideas and Outcomes* 3 (Apr. 5, 2017), e13086. ISSN: 2367-7163. DOI: 10.3897/rio.3.e13086.
- [7] J. Cardoso et al. “DCSO: Towards an Ontology for Machine-Actionable Data Management Plans”. In: *J Biomed Semant* 13.1 (July 26, 2022), p. 21. ISSN: 2041-1480. DOI: 10.1186/s13326-022-00274-4.
- [8] R. K. Singh and D. P. Madalli. “DMPFrame: A Conceptual Metadata Framework for Data Management Plans”. In: *Journal of Library Metadata* 23.3-4 (Oct. 2, 2023), pp. 121–160. ISSN: 1938-6389. DOI: 10.1080/19386389.2023.2268474.
- [9] T. Miksa, M. Jaoua, and G. Arfaoui. “Research Object Crates and Machine-actionable Data Management Plans”. In: (2020). DOI: 10.4126/FRL01-006423291.
- [10] E. Papadopoulou et al. “Data Management Plans as Linked Open Data: Exploiting ARGOS FAIR and Machine Actionable Outputs in the OpenAIRE Research Graph”. In: *Journal of Biomedical Semantics* 14.1 (Nov. 2, 2023), p. 17. ISSN: 2041-1480. DOI: 10.1186/s13326-023-00297-5.