Understanding the Data Management Plan as a Boundary Object Through a Multi-Stakeholder Perspective

Live Håndlykken Kvale
Oslo Metropolitan University

Nils Pharo
Oslo Metropolitan University

Abstract

A three-phase Delphi study was used to investigate an emerging community for research data management in Norway and their understanding and application of data management plans (DMPs). The findings reveal visions of what the DMP should be as well as different practice approaches, yet the stakeholders present common goals. This paper discusses the different perspectives on the DMP by applying Star and Griesemer's theory of boundary objects (Star and Griesemer, 1989). The debate on what the DMP is and the findings presented here are relevant to all research communities currently implementing DMP procedures and requirements. The current discussions about DMPs tend to be distant from the active researchers and limited to the needs of funders and institutions rather than to the usefulness for researchers. By analysing the DMP as a boundary object, plastic and adaptable yet with a robust identity, translating between worlds (Star and Griesemer, 1989) where collaboration on data sharing can take place, we expand the perspectives and include all stakeholders. An understanding of the DMP as a boundary object can shift the focus from shaping a DMP which fulfils funders' requirements to enabling collaboration on data management and sharing across domains using standardised forms.

International Journal of Digital Curation
2021, Vol. 16, Iss. 1, 19 pp.

1

http://dx.doi.org/10.2218/ijdc.v16i1.746
DOI: 10.2218/ijdc.v16i1.746

Introduction

The data management plan (DMP) is promoted as a tool for ensuring good data management and a first step for making data as open as possible, enhancing reproducibility and reusability of collected data and avoiding data loss (Michener, 2015). Funders increasingly require DMPs to be submitted along with research proposals and updated during the research projects (European Research Council, 2017b). Following funders requirements, a growing number of either discipline- or funder-specific templates and tools for DMPs have been developed. Within the research data management community the current focus is on making machine actionable, readable and interoperable DMPs exploiting the "thematic, machine-actionable richness with added value for all stakeholders" (Miksa et al., 2019). Several studies on DMPs take a quantitative approach to measuring effects, either in actual shared data or as successful funding (Diekema, Wesolek, and Walters, 2014; Johnson and Knuth, 2016; Mischo, Schlembach, and O'Donnell, 2014; Van Loon et al., 2017; Westra, 2017). Other articles focus on the importance of writing a DMP (Nature, 2018) or how to write one (Burnette, Williams, and Imker, 2016; Michener, 2015; Wright, 2016). All are useful and applied approaches with a focus on meeting requirements and receiving funding. This paper takes a different approach by aiming to understand the DMP as an object and document in the research process by investigating how the DMP is perceived by different stakeholders that all claim an interest in the plan and the planning.

The research questions investigated in this paper are:

1. What perspectives on DMPs are held by different stakeholder groups?

2. How do these perspectives help or hinder DMPs as tools to support data management?

Background

In 2017, the European Union's[1] (EU) Horizon 2020[2] (H2020) programme updated the EU's research data policies to require that new projects funded by the programme had to create a DMP (European Commission, 2016; European Research Council, 2017b). A policy document from the Norwegian Ministry of Education and Research (Kunnskapsdepartementet, 2017) made recommendations for universities and university colleges to require DMPs. The policy document described the DMP as a document containing plans for how research data will be managed through the research lifecycle to make data sharing an embedded part of the workflow. Further it should be a guiding document to help researchers in the project planning. Also the DMP should aid institutions, the research council and others in ensuring that the requirements are met, and it should serve as inspiration for other researchers to learn best practice (Kunnskapsdepartementet, 2017). The DMP is also expected to increase awareness and improve the way researchers document data and to enhance reproducibility (Kunnskapsdepartementet, 2017). The current template for H2020 DMPs from the

---

[1] Also funding associated countries, such as Norway, under the same conditions.
[2] Horizon 2020: https://ec.europa.eu/programmes/horizon2020/en/what-horizon-2020

European Research Council (ERC) focuses on how data can be made findable, accessible, interoperable and reusable (FAIR) and describes the costs associated with data management (European Research Council, 2017a). Unlike the EU (European Research Council, 2017a) guidelines, the policy document from the Norwegian Ministry of Education and Research emphasises that each institution is responsible for approving the DMPs (Kunnskapsdepartementet, 2017), and calls for the design of a tool to support development of DMPs. This has led to two national research data storage providers developing and publishing generic DMP tools[3].

In the EU and in Norway as addressed in this study, DMPs are relatively new to all stakeholders, including the researchers. The European Union ran a pilot requesting DMPs from selected thematic areas funded by Horizon 2020 between 2014 and 2016. In 2017, this pilot was extended to cover all areas of Horizon 2020 (European Commission, 2013; European Commission, n.d.). In Norway, a pilot on DMPs for climate research was done from 2014 to 2015, the aim being to learn whether DMPs would encourage more data sharing. This attempt was regarded as unsuccessful due to a lack of experience and knowledge for evaluating the DMPs amongst reviewers (Schjølberg, 2015). In its 2017 policy, the Norwegian Ministry of Education and Research points to research institutions as responsible for assessing DMPs. In doing so, they shift the focus from the evaluation of DMPs as part of funding applications to the creation of DMPs as part of research workflows. Consequently, universities are now establishing workflows for DMPs (NTNU, n.d.; The Artic University of Norway, 2019; University of Oslo, 2019). There are no national guidelines or criteria for evaluation of DMPs.

In the United States, DMPs have been a standard requirement in grant application for a decade (Mischo et al., 2014), and analysis of DMP guidelines and DMPs is an established part of the literature on data management (Berman, 2017; Burnette et al., 2016; Diekema et al., 2014; Dressel, 2017; Hardy, Hughes, Hulen, and Schwartz, 2016; Johnson and Knuth, 2016; Thoegersen, 2016; Van Loon et al., 2017; Williams et al., 2017; Wright, 2016). In reading and analysing the literature two tendencies become evident. Studies on DMPs tend to present the perspectives of one or two stakeholders and thus cover different aspects of and approaches to the DMP. Steinhart, Chen, Arguillas, Dietrich and Kramer (2012) use a survey to investigate researchers' experiences with DMPs. Researchers' perspectives are found in case studies describing how the DMP was applied in a research group (Burnette et al., 2016; Dressel, 2017) or discipline (Dressel, 2017). Other studies use quantitative approaches to measure the effect of DMPs either by grant success rates (Mischo et al., 2014) or by evaluating the effectiveness of research support by assessing the quality of DMPs (Johnson and Knuth, 2016; Van Loon et al., 2017). These studies present a research support perspective. Two studies have used content analysis to assess the requirements from the funders (Thoegersen, 2015; Williams et al., 2017). The results from these studies are useful for assistance in the writing of DMPs. Diekema and colleagues (2014) investigate researchers, research offices, and academic libraries in the role of infrastructure providers. They find that although researchers often are positive towards sharing data, they lack the necessary skills to do so. Researchers were unfamiliar with data repositories and existing data management services from the library. Further, they noted that data management mandates had little impact on the workflow of researchers and research office respondents. Diekema and colleagues propose that the library needs to make

---

[3] Examples include NSD (https://nsd.no/arkivering/en/data_management_plan.html) and Sigma2 (https://www.sigma2.no/content/easydmp)

researchers aware of existing research data management services and infrastructure to bridge the data management skills gap (Diekema et al., 2014).

Current literature on DMPs presents an applied approach to the document as a tool, and on how making the DMPs machine-actionable can be beneficial for multiple stakeholders (Cardoso, Proença, and Borbinha, 2020; Miksa et al., 2019; Simms et al., 2017). Less emphasis is placed on the content of the plan, why the plan is written and for whom. The review of DMP literature by Smale and colleagues (2020) does, however, suggest that there is no evidence to support a claim that researchers benefit from filling out a DMP. This suggests that it might be a good idea to take one step back and problematise the influence of the varying interests held by different stakeholders when creating a DMP and the tension between these interests.

Leading theorists in the area of data management emphasise that different stakeholders sometimes hold conflicting interests (Bowker, 2005) and that including multiple stakeholders when examining the functions of data in scholarship is important (Borgman, 2015). We believe this multi-stakeholder approach is the strength of the study presented in this paper, providing a broader understanding of the DMP, which is helpful in the practical approach to writing a DMP.

Theoretical Framework

The main characteristic of boundary objects is that they mean different things to groups of people working in different contexts and facilitate coordination and collaboration between these different groups. According to Star and Griesemer:

> 'Boundary objects are objects which are both plastic enough to adapt to local needs and the constraints of the several parties employing them, yet robust enough to maintain a common identity across sites. They are weakly structured in common use and become strongly structured in individual-site use. These objects may be abstract or concrete. They have different meanings in different social worlds but their structure is common enough to more than one world to make them recognisable, a means of translation' (Star and Griesemer, 1989).

In this paper, we will focus on standardised forms, described as methods of communication between different groups with different interests. One example of a standardised form which Star and Griesemer use in their original work on boundary objects is a document of procedures for data collection and curation, 'a precise set of procedures for collecting and curating specimens' (Star and Griesemer, 1989). This description of a precise set of procedures for collection has strong similarities with some of the descriptions we found of the DMP. Still, Star acknowledges the challenges in collecting, disciplining and coordinating distributed knowledge (Star, 2010). Her example of the complexity of creating forms is from a research study on epileptic patients from late nineteenth century England. She found that much information was scribbled down on the edge of the form because it did 'not fit the actual form' (Star, 2010). The information was later discarded as unimportant because it was not part of the information family members of the epileptic patients were asked to collect. Star describes these documents as "revealing the relations of class and medicine" in England at the time. Thus, Star asks, "how do forms shape and squeeze out what can be known and collected?" (Star, 2010). This is a problem which emerges frequently in the era of

automation and digital forms: there is often little space for scribbling on the side. The different approaches to the standardised form, either as a set of procedures or boxes to fill in supplying the requested information, call for different levels of involvement from the contributor.

Method

We have used a modified Delphi study (Ziglio, 1996) to explore the understanding and application of DMPs among different Norwegian stakeholders involved in research data sharing. A Delphi study is characterised by the use of an expert panel to elicit opinions on a shared reality from different perspectives. Data collection is performed in several rounds with the intention of reaching consensus or solving an issue.

A group of 24 experts took part in the study. Table 1 contains an overview of the participants. The group consisted of policy-makers, representatives of national service providers, and researchers and research support staff from four Norwegian universities. The participants were invited based on their involvement in the development of policies, infrastructure or data-related research support. The research support staff were recruited to include representatives from different research support services at the universities, including libraries, research offices and IT departments.

We invited researchers who were appointed as project owners of H2020-funded projects to participate in the panel. Of the 25 researchers contacted, eight participated. These eight researchers hailed from different disciplinary backgrounds (biology, musicology, science studies, economics, neuroscience, psychology, philosophy, gender studies), and they differ in levels of prior knowledge of research data management.

The participants were not promised anonymity, only that their names would not be used. Identification might be possible with triangulation and local knowledge. As a result, quotes in cases in which the informant does not wish to be identified in connection with the statement do not include the full participant code.

Table 1. The participants organised according to role.

Role/Stakeholder Category	Participant Code			
Researchers working individually	RIZ	RIJ	RIL	RIB
Researchers working in groups	RGV	RGD	RGA	RGW
Policy-makers	POU	POS	POK	
Service providers	INH	INO	INR	
Research support IT	ITE	ITY	ITI	
Research support, research office	ROC	ROX	ROT	
Research support, library	LM	LP	LG	LN

Data were collected in three phases, as shown in Figure 1. The first phase, the 'exploration phase', was conducted using open interviews lasting approximately one hour in January/February 2018. The purpose of this phase was to obtain an initial overview of the panel members' opinions on the DMP, or 'defining the problem'. Interviewees were asked a set of questions concerning research data management,

including their needs for data management, their experiences with DMPs and their perceptions of the aim of a DMP.

In the second phase, the 'evaluation phase', conducted in August/September 2018, participants answered a survey containing nine questions on topics such as data stewardship, DMPs, ethical aspects of data sharing and core functions in a research data infrastructure. The survey was designed to further explore issues and tensions uncovered in the first interviews. Several of the questions were formulated as statements that the participants were asked to agree or disagree upon.

The third, 'concluding phase', was conducted using interviews in March/April 2019. These interviews lasted approximately 30 minutes and were based on results derived from the two former phases. Among the questions asked in the final interview were: How does the DMP best reflect the different needs of the different stakeholders, and participants were asked whether they had thoughts on the preliminary findings of this study, such as the differences reported by researchers working individually and in groups.

Based on requests from some of the participants, the questions were sent to all participants prior to the data collection, in all three phases. The participants were also sent the transcripts from the interviews and were asked for permission to share the complete material or parts of the data material to which they contributed. The data is available in Zenodo (Kvale, 2020). In this paper, data regarding DMPs from all three phases are reported and integrated in the analyses (Creswell et al., 2018). The interviews were qualitatively coded and analysed thematically (Saldaña, 2016).
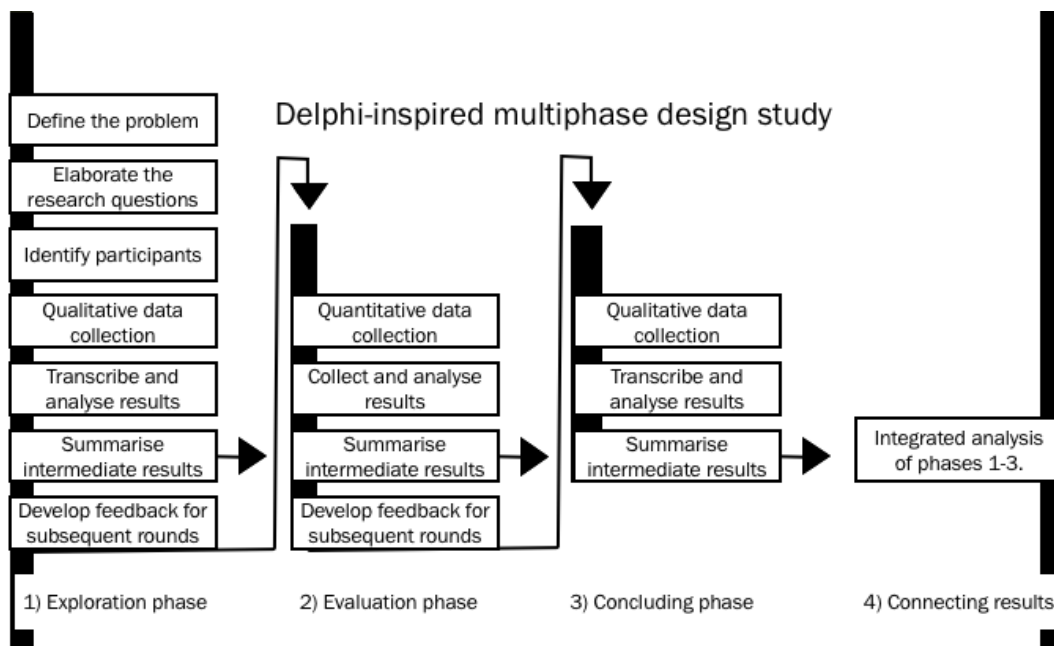

Figure 1. A Delphi-inspired multiphase design study.

Findings

The findings reported here are based on the integrated analysis of the material from all three phases of the study. Findings were subsequently grouped according to three main

themes to highlight different issues regarding the DMP: Sharing a common goal, Different perspectives on the DMP, and Different practice approaches to the DMP.

Sharing a Common Goal

Analysis of the first interviews revealed 12 different perceived purposes for using a DMP amongst our participants (Table 2; rows A-L). These purposes were used in the survey to understand the extent to which there was agreement among stakeholders about the purpose of a DMP. To cover other views expressed in the exploration phase, three additional options were added (rows M-O). In the survey, the participants were presented with a list of purposes for making a DMP and were asked to select the five most important reasons to make a DMP. The third column in Table 2 shows the number of times each of the aims was selected.

Table 2. Aims of the DMP. (n = 24 participants).

	Reasons to make a DMP	Frequency
G	It makes the researcher think about how to make their data and metadata FAIR (findable, accessible, interoperable, reusable).	21
H	A DMP creates awareness and agreement on data procedures within a research group, which is particularly important when several partners with different data needs are involved.	19
B	When researchers are in control of their own data management, the research gets better and more efficient.	15
D	When researchers think about what they are doing at an early stage, they can make intelligent choices for their data.	15
K	A DMP is a plan with a commitment to making the data as open as possible.	6
C	A DMP gives universities an overview of ongoing research projects.	5
A	To create awareness in the research community of the need for data stewardship.	5
L	A DMP makes researchers aware of the value of what they are collecting.	4
F	A DMP gives the archive the possibility to plan for data which are going to be deposited there.	3
O	I am not familiar with data management plans.	3
E	A DMP provides the archives with information (metadata) they need when data is deposited.	3
N	I don't see why DMPs are important.	2
M	Other reasons: 'Power and competency to avoid	1

	Reasons to make a DMP	Frequency
G	It makes the researcher think about how to make their data and metadata FAIR (findable, accessible, interoperable, reusable).	21
	ethical brakes in terms of personal privacy'.	
J	A DMP is primarily a tool for those who take care of the data after it has left the project.	1
I	A DMP shows what data will be produced in a project when the project is finished.	1

Four aims (G, H, D and B) were selected significantly more often than the others (by 15 to 21 of the participants).

The most important reasons to make a data management plan (as selected by participants) are:

- **G:** The DMP makes the researcher think about how to make their data and metadata FAIR (findable, accessible, interoperable, reusable).

- **H:** A DMP creates awareness and agreement on data procedures within a research group, which is particularly important when several partners with different data needs are involved.

- **D:** When researchers think about what they are doing at an early stage, they can make intelligent choices for their data.

- **B:** When researchers are in control of their own data management, the research gets better and more efficient.

All of these reasons emphasise the researcher both as the creator of the plan and the primary beneficiary of thorough planning. The different stakeholders agreed on a common goal of a data management plan. Aims G, H, D and B have in common that the goal of a DMP is to improve data management by making researchers plan for sharing their data internally within research groups and externally (FAIR) by creating procedures for documentation and collection at an early stage. The survey brings the areas of agreement to the surface, and it therefore appears to be a broad agreement among different groups of stakeholders about the purpose and role of DMPs. The interviews, however, tell a different story, with perspectives and approaches varying according to the different contexts in which each group of stakeholders work.

Different Perspectives on the DMP

The first interviews reveal five different perspectives on the DMP reflected by the vision of the stakeholders and have been analysed, grouped and labelled accordingly.

The participants representing policy-makers and research support services agreed largely on the DMP being a reflection of the extent to which data can be shared and on how data sharing is an aspect of open research. One of the librarians stated "it is about the researchers already in the design phase reflecting on how to work as openly as possible" (LG). Meanwhile, one of the policy-makers focussed on the management, publication and associated costs for which the DMP should be used to prepare: "What type of data to collect and how to take care of them, how to make them available and

possibly how to fund data management" (POU). Another policy-maker focused on similar aspects by putting forward the need for data stewardship: "It is for the whole research environment to become aware of their need for data stewardship" (POK). Both policy-makers hold a funder perspective on the DMP, emphasising that it is used to manage how data can be made available and enable the calculation of data sharing costs. The librarian, on the other hand, focussed more on the structured planning for the research process with data sharing as the ultimate goal. We have labelled these *the sharing and open science perspective* and *the stewardship perspective*, respectively.

Researchers have divergent views on what the DMP is, based on whether they work in collaborative environments or in more individual-based research environments in which the sharing of data among collaborators is less common. RGV, RGW, and RGD are all researchers who work in groups in which data is shared both within the group and with external partners. They described what is categorised as an internal *protocol and procedures perspective* on the DMP. In contrast, the individual-oriented researchers (RIZ, RIB, and RIL) had no experience with DMPs. Both RGW and RGV described the DMP as a document used for agreeing on standard procedures. As such, the DMP becomes more of an internal document for the respective research group. Researcher RGD described the entire research project as a DMP: "Actually, the whole project is like a big data management plan" (RGD). The research project RGD is referring to combines data from different locations and previous research in a new databank for the researchers to collaborate on. As it is described, the research project itself is to a large extent about managing data, and the description of the project becomes the DMP. In the second interview, one year later, RGD described another DMP document used in the same project. In it, the data manager had created a detailed protocol for how to work with the data in the project to ensure that all researchers involved in the project followed the same procedures when working on the existing data or adding new data.

Researchers working independently, or in collaborations in which there is little or no sharing of data among collaborators, express less knowledge of the DMP. Researchers RIZ, RIB, and RIL were, as stated above, unfamiliar with DMPs. However, RIB was familiar with aspects of data management and shared comprehensive descriptions of how she analysed data and how the data could be accessed as supplements to journal articles. Documenting data was a clear part of RIB's research, even if there were restrictions on sharing the data. Researcher RIJ is a researcher within philosophy with experience of ethical committees and a strong interest in privacy protection and research ethics. RIJ's understanding of a DMP was similar to that of the policy-makers and research support staff, with more emphasis on aspects regarding personal privacy.

The service providers presented a more differentiated view. INR focussed on how the DMP is useful for several stakeholders, stating it is "a tool for planning with archiving and sharing in mind" (INR) and, "for the researcher and the institutions to make sure their researchers fulfil the demands" (INR). This aspect of control for the institution was not promoted by other stakeholders in this study. INO focussed on decisions that should be made prior to data collection: "for the researcher to think about what he is doing at an early stage, so he can make intelligent choices" (INO). INH emphasised that the DMP is a document the researchers do not create themselves: "When I got this task [to create a DMP tool], I thought those that are using my tool are going to be researchers, but although we put all the effort to facilitate the creation of a DMP, it still requires some competence in data management that is not likely to be present in the end user, in the researcher" (INH). These quotes from the service providers present a *curating perspective* and *fulfilling requirements perspective*.

Stakeholders view of users

The survey asked participants to name who the DMP primarily is written for and rank the users of the DMP according to their importance on a scale of 1-3. The results show that the participants, to a large degree, agree on the DMP being for the researcher (a score of 52 out of 72). Other central users of the DMP are the research institutions (24/72) and other researchers (17/72). In addition, funders (9/72), archives (5/72) and data stewards (2/72) were mentioned. To illustrate the different opinions, a relational visualization is used to show relationships and connections between the data (Figure 3). This visualization shows that the different stakeholder groups point at different users of the DMP, and that there is no clear coherence in the responses apart from the common agreement about the researcher. All participants point at several users of the DMP, which again illustrates the different perspectives of the plan listed above. Accordingly, DMPs should be developed to be used by different stakeholders for different purposes, with primary focus on the researcher. To follow up on this aspect we invited the participants to give concrete suggestions on practice approaches to how the plan could be perceived as useful for the different stakeholders.

Different Practice Approaches to the DMP

In the final interviews, the participants were asked to make suggestions about how the DMP could be developed to respond to the needs of its different users, and specifically about how the DMP should be developed to become a useful document. The different stakeholder groups suggested different approaches to developing the DMP as a practical tool.

Among the research support services, participants emphasised the DMP as a document in which to display and encourage best practices on data management by embedding checklists and good examples. The document should, according to them, be developed to reflect the researchers' perspective on the research process. One of the participants working in IT support put it like this:

> 'You should give some kind of best practice, both tips and strategies, for how you as a researcher should do best practice on data. If you ask questions that not just irritate the researcher but rather enlighten them on something they didn't know. The questions should make them think "wow, I do have to think about this", I believe that this could be a useful approach' (ITI).
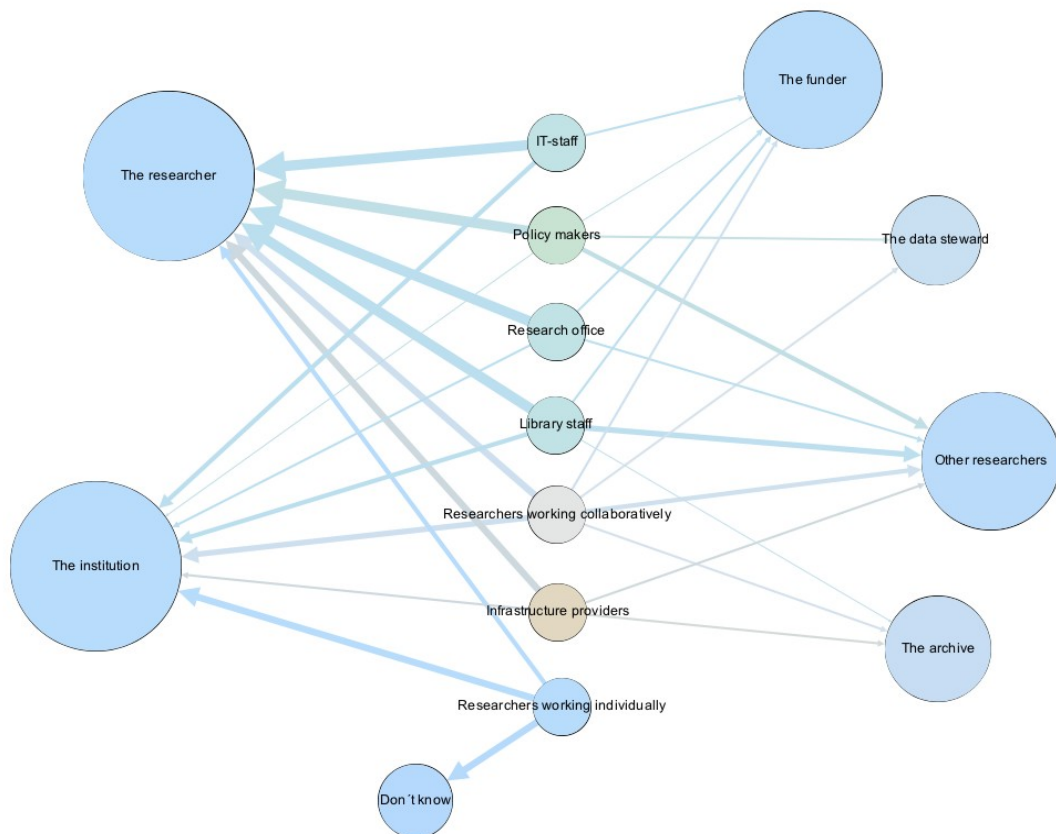
Figure 2. Relational visualization of responses to the question: Who is the data management plan primarily for?

One of the library staff participants described it in a similar way:

> 'What I think and believe will be important is informative help texts for the different sections, in a way translating the computer syntax, that you might have a bullet point checklist or question that the researchers should ask themselves when answering that section' (LM).

Aiming at creating guidance that encourages researchers to reflect on data management practices was typical among the research support staff. One participant was concerned about how the interests of the other stakeholders could reduce the plans' relevance and usefulness for the researchers:

> 'We see in Norway that there are some entities who think they should use the DMP for all other types of purposes, to their own advantage. And it is possible that there will be types of secondary use, for the institutions to monitor research, and for archives it might make the archiving process easier, but that should not be the primary function, and one should not create templates focussing on this instead of the researchers' needs' (LG).

Their concern was that the interests of other stakeholders will make the DMP less useful for researchers. The research support staff therefore suggested the creation of a list of questions formulated in such a way that researchers find it useful to reflect upon

the questions, supplemented with tips and best practices on how to improve data management.

The researchers were looking to their international research communities for best practice and evolving methods, standardisation and expectations. This was expressed both by RIB and RGW:

> 'It is evolving. Some groups have high data management standards and you try to adapt to it. But that is also expensive, depending on what you do, so then if that is the standard, we have to invest in that kind of resources. So, specifically, in the data we are influencing each other internationally' (RGW).

> 'Within many areas there has become an increased focus on reproducibility and then it is important to actually have the data, the codes and what else might be needed for replication' (RIB).

The increased awareness described by RIB was found among several stakeholders in the final interviews and will be investigated separately in a forthcoming paper. RGV described the need for simple language and relevance.

> 'I realize that some believe that the current forms [referring to experience with existing templates] are fairly simple and clear; at the same time it is in practice very difficult' (RGV).

RGV further presents the idea of using a decision tree to visualise and decide what is relevant for different researchers to consider in their plans and to supply the research office with extensive knowledge of data management to guide the researchers in writing the plan.

> 'I think what really matters, when these in reality often are complex issues, is another person between us and the [service provider] ….and the person managing this as an advisor must have extensive knowledge of the whole field, not just disciplinary glasses on, seeing only what is relevant in medicine or sociology, but one that understands the background for certain questions and understands what is important to maintain and legitimate interests for those who are part of the research' (RGV).

This request points at the data steward role, i.e. someone somewhere in the university with competence in data management.

One of the researchers described a lack of coherence between policy and practice:

> 'Basically, I had a chat with EU and they say that once your DMP is accepted, unless it is absolutely necessary, please do not spend too much time on it because the main goal is the scientific research, and we have only two years of funding, so then, yeah, so we keep it as such but there are not big changes in the way I manage my data, so I did not really think about it again' (Researcher).

This reflects notions of a funder with no clear interest in data management, revealing a conflict between the requirements for updating a DMP and the standard model (Bowker, 2005) for scientific publication.

Among the policy-makers, there were different opinions on the DMP. However, they do agree that it should be a useful document for researchers. POK was clear in her opinion that, 'it should not be up to the authorities to specify what type of DMPs are good for the researchers to make; that is none of our business' (POK). Her point was that the researchers should write the DMP in a manner they find it useful rather than being forced to use a certain template or tool to create it. POU expressed concern about the extent to which the DMP is relevant for researchers in its current form:

> 'Today the DMP is formulated a bit like a questionnaire, generating a PDF and that's the way it gets, maybe not that useful, it becomes more like an exercise, a bureaucratic exercise' (POU).

She further explains how it is difficult to complete, 'Because concepts like metadata are not something most researchers relate to' (POU). She also proposed automating parts of the DMP creation process,

> '...so that some information can be automatically added, and others be automatically proposed. Ideally, the calculation of data management costs can increasingly be automated' (POU).

What POU proposed is automating as much of the DMP as possible so that researchers do not have to spend time on trivial questions. Her approach stands in contrast to that of both POK, who wished to minimise the formal requirements of a DMP, and POS, who described the function of the DMP in a way similar to that of the research support services.

> 'What I communicate to my researchers is that you will always generate data. Describing it in a DMP, even if you work on an individual research project, and being explicit about how you are going to structure your data, so that you actually can reproduce your results at a later stage, improves the research process [….] I don't know if this is currently reflected in the DMP, but I believe it is a way to strengthen the research process' (POS).

Among the service providers, we found a change in their views of the DMP compared with what they had expressed in the earlier phases of this study. One service provider, a technician and service provider delivering tools to the universities, referred to the DMP as a complex document serving several purposes and stakeholders:

> 'I have only become more uncertain about what a DMP is; the more I try to understand it, I am certain that I had a simpler perception of the problems last time we spoke, and I am a bit frustrated over my own lacking capability to get anywhere. Because it is important to very many, but for different reasons something that results in it being perceived as pretty useless for everyone because it tries to solve too many things at once' (INO).

This statement by INO reflects some of her difficulty in understanding research. In a prior stage of the study INO focussed on the institutions and their wish to have an

overview of research data collected and control over where and how it was archived, but in the year which had passed she had become more uncertain about the DMP.

Another service provider, INH, emphasised meeting the needs of the researchers. In the first interview, INH did not believe that the researchers would be capable of filling out the DMP. In the later interview, however, she stated that:

> 'Researchers have to upgrade along the way in the research process, and this as the DMP should be drafted or ready since the very beginning because it is part of, I mean it is part of the research process itself. So, making the plan is not for the sake of making a plan; it is part of the research' (INH).

INH also suggested that the DMP should be developed to serve the needs of the researchers:

> 'First of all, the researchers have to interact actively with the DMP, so it has to be in electronic form, it has to be modified and customised as much as possible, so this means that actually the guidelines should be really high-level guidelines, but the action form of the DMP should be dedicated to the community specific. Only in this way can you make sure that it is not an [exercise of] checking boxes' (INH).

INH further emphasized that data management should be "customized to the scientific topic" and a digital tool.

The understanding presented by INR overlaps with that of INH, who also struggled to develop a relevant DMP tool:

> 'It should not just become a questionnaire ending in a document you send to the funder because they require it' (INR).

Still, her approach to guiding researchers in the right direction is somewhat different from that of INH, who focused on a community specific DMP, whereas INR focussed on embedding a detailed level of institutional guidance:

> 'There should be a guidance in the DMP so that when you answer questions you are guided in the right direction. So that with naming conventions, really what to name the files, and how to structure data, there might be similarities, and then you can get help and suggestions as to how you should name your files' (INR).

Another suggestion by INR was to use a guide for the classification of data according to sensitivity, so that, while writing the DMP, the level of sensitivity is defined for the data to be collected.

The different practical approaches to DMPs point in different directions. The researchers requested on-the-ground support, and development of a peer-network to share best practices. The research support staff focussed on well-formulated questions encouraging researchers to reflect and make decisions for data management. Both approaches imply a lower level of automation and a higher level of flexibility or a more manual plan. Among the policy-makers, the opinions differed: one was clear that such decisions should be left to research communities, another focussed on the importance of the DMP as a useful tool for researchers, whereas the third suggested that more

information should be automatically added, a notion shared by the infrastructure developers.

Discussion

The findings presented perceptions of the DMP held by the different stakeholders. The different stakeholders understand and apply the function of the DMP differently: the curating and fulfilling requirements perspective, the sharing and open science perspective, the stewardship perspective, and the protocol and procedures perspective. The perspectives illustrate the different backgrounds of the various stakeholders. The researchers reflect on how DMPs could be useful in a research group, while the service providers reflect on how they can be used to assist in planning for the archiving process and fulfilling formal requirements. In Star and Griesemer's terms, this would constitute worlds of the different actors. Research, in general, and the sharing of research data, in particular, requires cooperation between different stakeholders. By writing a DMP, researchers plan for their data to move from collection through analysis and to sharing as was agreed upon in the goal of the DMP. In this sense, DMP creation facilitates translation between the different worlds and the different stakeholders as a standardised form by creating context for research data so that these can be understood and interpreted in the different worlds independently of disciplines, institutions or national boundaries.

The ideal DMP, as described by the support services participants, is a guiding tool which poses questions researchers have not thought about. The service providers presented a different understanding, focussing on meeting formal requirements and possibilities of automation. The researchers look to their international research communities for best practice and request support functions in their research institutions. However, in encounters with policy-makers, researchers are confronted with the blunt reality of little time and money to think about data quality and the continuous pressure to publish. The tension between different perspectives, the research reality and the higher goals, can be resolved by a common understanding of the DMP as a document which is not a product of consensus, but an everyday translation between worlds and communities. Star (2010) points at a common misunderstanding regarding the need for consensus as a basis for cooperation, claiming that the use of boundary objects can explain how cooperation can continue unproblematically without consensus. The DMP may perform the role of a boundary object for different data management stakeholders.

We find that there are two issues which need to be clarified in order for the DMP to function optimally as a standardised form translating between worlds, formalising procedures and standardising methods: the *degree of standardisation* and the *degree of automation*.

When it comes to degree of standardisation, the policy-makers problematised how the DMP today becomes more of a bureaucratic exercise than an actual plan and emphasised that they do not want to interfere with what should be in the plan. At the same time research communities are continuously developing best practices for data management and there is no static standard for how data management should be done. This suggests that a lower degree of standardization would give the DMP flexibility over time and across methods and disciplines. The DMPs should therefore be developed more as open documents to fill the needs of the researchers, in their planning for sharing of the data. The DMP is never a goal in itself, rather it is the reflection it triggers

regarding data sharing that is the desired output. The different stakeholders agree on the goal of a DMP: to improve data management by making researchers plan for sharing their data internally within research groups and externally (FAIR) by creating procedures for documentation and collection at an early stage.

To achieve this, research support and infrastructure developers need to take one step to the side and leave the researchers with autonomy to shape the content of DMPs according to the design of their research projects. We suggest formulating open-ended questions concerning data management issues for researchers to reflect upon how data best can be structured and documented for reuse and sharing. In addition, researchers should be supported with best practices to ensure high-quality data management.

The degree of automation refers to tools for DMPs and how they should be developed. Information which is on a general level or project information could preferably be imported or connected to other sources. Harvesting data from DMPs to repositories and research administrations tools does, however, come at the cost of the autonomy of a plan. Automated decisions do require a preselection of options, which again would be limiting the possible choices for the researcher. We therefore argue that the level of detail in the DMP itself will and should vary significantly between research projects. Automated input of general information could be useful, this type of information should however be kept to a minimum. Further, we do not find that automated decision-making and harvesting standardized output is beneficial for the researchers in their planning of data management.

Our suggestion is therefore to focus on balancing the guidance and decision-making, leaving flexibility for the researcher in the creating the DMP.

Conclusion

Creating consensus between data management stakeholders might not be necessary for cooperation or successful conduct of data management planning. With respect to our first research question, we identified four different perspectives amongst the participants. The different stakeholders have different perspectives each of which reflects to some extent their backgrounds and roles. The perspectives need to be considered if the DMP is to work as a document translating between different stakeholders and supporting the longevity of research data. Despite conflicting approaches to how the DMP should be developed, the stakeholders agreed on a common goal of creating the DMP and that the DMP has a purpose for several stakeholders, including themselves. Considering our second research question, our findings suggest that conflicting perspectives currently result in researchers becoming more distanced from the DMP, and that DMPs risk becoming merely a bureaucratic exercise. If leaving the shaping of the plan more open to the researchers to adapt to their needs, it can become useful in helping researchers plan for data sharing. The DMP should allow researchers to scribble down what is most relevant in each unique research project. The lack of coherence and the complexity of DMPs could be turned into a strength. If the DMP is to function as a standardised form facilitating co-ordination and collaboration between different groups of people, the degrees of standardisation and automation must be balanced, leaving the researchers with flexibility in the development and implementation of the plan. Only then can the DMP function as a boundary object translating between worlds. By formalising procedures and standardising methods, the DMP can become a boundary object, enhancing reproducibility and enabling data sharing.

References

Berman, E. (2017). An exploratory sequential mixed methods approach to understanding researchers' data management practices at UVM: Integrated findings to develop research data services. *Journal of eScience Librarianship, 6*(1), e1104. doi:10.7191/jeslib.2017.1104

Borgman, C. L. (2015). Big Data, Little Data, No Data: Scholarship in the Networked World. Cambridge, MA: MIT Press.

Bowker, G. C. (2005). Memory Practices in the Sciences. Cambridge, MA: MIT Press.

Burnette, M., Williams, S., & Imker, H. (2016). From plan to action: Successful data management plan implementation in a multidisciplinary project. *Journal of eScience Librarianship, 5*(1), e1101. doi:10.7191/jeslib.2016.1101

Cardoso, J., Proença, D., & Borbinha, J. (2020). Machine-Actionable Data Management Plans: A Knowledge Retrieval Approach to Automate the Assessment of Funders' Requirements. In J. M. Jose, E. Yilmaz, J. Magalhães, P. Castells, N. Ferro, M. J. Silva, & F. Martins (Eds.), Advances in Information Retrieval (pp. 118–125). Springer International Publishing. doi:10.1007/978-3-030-45442-5_15

Creswell, J. W., & Plano Clark, V. L. (2018). Designing and Conducting Mixed Methods Research (3rd ed.). SAGE Publishing.

Diekema, A. R., Wesolek, A., & Walters, C. D. (2014). The NSF/NIH effect: Surveying the effect of data management requirements on faculty, sponsored programs, and institutional repositories. *Journal of Academic Librarianship, 40*(3–4), 322–331. doi:10.1016/j.acalib.2014.04.010

Dressel, W. (2017). Research data management instruction for digital humanities. *Journal of eScience Librarianship, 6*(2), e1115. doi:10.7191/jeslib.2017.1115

European Commission. (2013). Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020. Retrieved from http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

European Commission. (2016). Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020. European Commission.

European Commission. (n.d.). Data management − H2020 Online Manual. Retrieved from https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm

European Research Council. (2017a, April 12). European Research Council Data Management Plan template. Retrieved from ERC: European Research Council website: https://erc.europa.eu/content/erc-data-management-plan-template

European Research Council. (2017b, April 21). Guidelines on Implementation of Open Access to Scientific Publications and Research Data – In projects supported by the European Research Council under Horizon 2020. Retrieved from http://ec.europa.eu/research/participants/data/ref/h2020/other/hi/oa-pilot/h2020-hi-erc-oa-guide_en.pdf

Hardy, L. J., Hughes, A., Hulen, E., & Schwartz, A. L. (2016). Implementing qualitative data management plans to ensure ethical standards in multi-partner centers. *Journal of Empirical Research on Human Research Ethics, 11*(2), 191–198. doi:10.1177/1556264616636233

Johnson, A., & Knuth, S. (2016). Data management plan requirements for campus grant competitions: Opportunities for research data services assessment and outreach. *Journal of EScience Librarianship, 5*(1), e1089. doi:10.7191/jeslib.2016.1089

Kunnskapsdepartementet. (2017). Nasjonal strategi for tilgjengeliggjøring og deling av forskningsdata (p. 48). Oslo: Kunnskapsdepartementet.

Kvale, L. (2020). Data from a three-phase Delphi study was used to investigate an emerging community for research data management in Norway: KIRDN_Data. doi:10.5281/zenodo.3673053

Michener, W. K. (2015). Ten simple rules for creating a good data management plan. *PLOS Computational Biology, 11*(10). doi:10.1371/journal.pcbi.1004525

Miksa, T., Simms, S., Mietchen, D., & Jones, S. (2019). Ten principles for machine-actionable data management plans. *PLOS Computational Biology, 15*(3), e1006750. doi:10.1371/journal.pcbi.1006750

Mischo, W., Schlembach, M., & O'Donnell, M. (2014). An analysis of data management plans in University of Illinois National Science Foundation grant proposals. *Journal of EScience Librarianship, 3*(1). doi:10.7191/jeslib.2014.1060

Nature. (2018). Everyone needs a data-management plan. *Nature, 555*, 286. doi:10.1038/d41586-018-03065-z

NTNU. (n.d.). Open Data. Retrieved from https://innsida.ntnu.no/wiki/-/wiki/Norsk/NTNU+Open+Data

Saldaña, J. (2016). The coding manual for qualitative researchers (3rd ed.). London: SAGE Publishing.

Schjølberg, R. (2015, May). Research Council of Norway's policy for open access to research data, data management plans. Presented at the NorStore Research Data Management training seminar, Trondheim, Norway. Retrieved from https://www.sigma2.no/sites/default/files/media/norstore_rune_nfr_-_data_retningslinjer260515.pdf

Simms, S., Jones, S., Mietchen, D., & Miksa, T. (2017). Machine-actionable data management plans (maDMPs). *Research Ideas and Outcomes, 3,* e13086. doi:10.3897/rio.3.e13086

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

Star, S. L. (2010). This is not a boundary object: Reflections on the origin of a concept. *Science, Technology, & Human Values, 35*(5), 601–617.

Star, S. L., & Griesemer, J. R. (1989). Institutional ecology, "translations" and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science, 19*(3), 387–420. doi:10.1177/030631289019003001

Steinhart, G., Chen, E., Arguillas, F., Dietrich, D., & Kramer, S. (2012). Prepared to plan? A snapshot of researcher readiness to address data management planning requirements. *Journal of EScience Librarianship, 1*(2). doi:10.7191/jeslib.2012.1008

The Artic University of Norway. (2019). Retningslinjer for oppfølging av datahåndteringsplaner ved UiTs enheter. Retrieved from https://uit.no/Content/633034/cache=20191706140204/Retningslinjer%20oppfolging%20DMP_publisert.PDF

Thoegersen, J. (2015). Examination of federal data management plan guidelines. *Journal of EScience Librarianship, 4*(1). doi:10.7191/jeslib.2015.1072

University of Oslo. (2019). Datahåndteringsplan ved UiO – Kompetanse, maler og organisering (p. 11). Retrieved from University of Oslo website: https://www.uio.no/for-ansatte/arbeidsstotte/fa/forskningsdata/291119_dhp_rapport_endelig-nettversjon.pdf

Van Loon, J. E., Akers, K. G., Hudson, C., & Sarkozy, A. (2017). Quality evaluation of data management plans at a research university. *IFLA Journal, 43*(1), 98–104. doi:10.1177/0340035216682041

Westra, B. (2017). Assessing data management practices and needs through data management plans. Abstracts of Papers of the American Chemical Society, 253.

Williams, M., Bagwell, J., & Zozus, M. N. (2017). Data management plans: The missing perspective. *Journal of Biomedical Informatics, 71*, 130–142. doi:10.1016/j.jbi.2017.05.004

Wright, A. (2016). Electronic resources for developing data management skills and data management plans. *Journal of Electronic Resources in Medical Libraries, 13*(1), 43–48. doi:10.1080/15424065.2016.1146640

Ziglio, E. (1996). The Delphi Method and its contribution to decision-making. In M. Adler & E. Ziglio (Eds.), Gazing into the Oracle − The Delphi Method and its Application to Social Policy and Public Health (pp. 3–33). London: Jessica Kingsley