Aligning Data Management Plans with Community Standards using FAIR Implementation Profiles

Navroop K. Singh[1][0000−0001−9131−3528], Shuai Wang[1][0000−0002−1261−9930], Angelica Maineri[2][0000−0002−6978−5278], Tycho Hofstra[3][0009−0009−7320−864X], Mark Bruyneel[3][0000−0002−9426−3439], Stephanie van de Sandt[3][0000−0002−9576−1974], Ronald Siebes[1][0000−0001−8772−7904], Jacco van Ossenbruggen[1][0000−0002−7748−4715], and Tobias Kuhn[1][0000−0002−1267−0234]

[1] Department of Computer Science, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands
`{shuai.wang|r.m.siebes|jacco.van.ossenbruggen|t.kuhn} @vu.nl,`
`n.k2.singh@student.vu.nl`
[2] ODISSEI, Erasmus School of Social and Behavioral Sciences, Erasmus University Rotterdam, 3000 DR Rotterdam, the Netherlands
`angelica@odissei-data.nl`
[3] University Library, Vrije Universiteit Amsterdam, De Boelelaan 1105, Amsterdam, the Netherlands
`{t.m.hofstra|m.bruyneel|s.van.de.sandt}@vu.nl`

Abstract. Data Management Plans (DMPs) are often required by organizations and funding agencies. When writing DMPs, taking into account community standards for publishing and managing research data can be a challenge for researchers. Community standards are often documented informally or communicated by word of mouth. The introduction of FAIR Implementation Profiles (FIPs) offers a structured way to capture such standards. This paper investigates with a user study, whether FIPs can serve as suggestions for aligning research data management with community standards. Through a customized interface with the related information extracted from FIPs as suggestions, we study whether researchers can take such suggestions into account when writing DMPs. Subsequently, a survey was conducted for each participant. The survey results highlight some specific DMP questions where FIPs can be effectively used as suggestions. The findings of the survey affirm the potential of FIPs as a valuable resource to harmonize research data with community standards.

Keywords: FAIR Implementation Profile · FAIR Principles · Data Management Plans

1 Introduction

To make research data findable and readily reusable by others, researchers are often mandated by funding organizations and universities to create *Data Management Plans* (DMPs). DMPs are generated from the corresponding templates

(a collection of related questions). When completed, DMPs are formal documents consisting of answers to questions that outline how data is handled throughout and after a research project. Answering these questions, researchers specify the details and methods of data collection, data repositories, responsibility, accessibility, and licences. These answers can be influenced by many factors: the requirements of conferences and publishers, the recommendations by departments and universities, the suggestions by colleagues, the community standards, etc. The following of community standards is a requirement of the FAIR principles (principle R1.3).[4] When uncertain about community standards, many choose to consult data stewards and colleagues. However, not all data stewards and colleagues are aware of the standards of every community, not to mention that this can be time-consuming. Moreover, community standards are often in the word of mouth or informally documented (and often inaccessible beyond the community/organization). Even worse, community standards can evolve as members of the community adopt new tools, repositories, registries, and licences. For this reason, aligning DMPs with community standards remains an unsolved problem.

The introduction of *FAIR Implementation Profiles* (FIPs) offers a structured way to capture community standards [8]. FIPs serve as structured templates about decisions and guidelines by experts and members of communities of practice [8]. Moreover, FIP comes with related tools and metrics that make comparison of community standards and statistical analysis easy [8]. The alignment of DMPs with community standards captured by FIPs has the potential to offer substantial benefits to both researchers and their respective communities. For example, this alignment would make data findable in uniform repositories and promote the standardization of some machine-interpretable format, which makes it easier to integrate it into a web framework and automatically compared other schemas. However, the realization of effective alignment faces several obstacles. The DMP templates universities/institutes use can vary significantly from each other. Some can have multiple versions based on the faculty and funding agency. Determining which questions in the DMP template could align with specific community standards can be ambiguous. Moreover, the datasets could be of interest to multiple communities, which further complicates researchers' efforts when selected for reuse. In this paper, we explore a convenient, reliable, and easy-to-use means for researchers to align with community standards using FIPs.

The idea of using FIPs as suggestions for DMPs was initially proposed by K. Hettne et al. [3]. However, they did not conduct any user study to validate this idea. In this paper, we take an empirical approach and explore the workflow to extract information from FIPs as suggestions on the DMP interface. We evaluate its efficacy with a survey. We aim to answer the following research questions.

- **RQ1:** Which questions in the DMP template can take community standards in FIPs as suggestions?
- **RQ2:** How can we build a user interface that takes community standards as suggestions?

---

[4] More details on the R1.3 principle are available at `https://www.go-fair.org/fair-principles/`.

– **RQ3:** How do users take advantage of suggestions from FIPs while writing their DMPs?

For RQ1, we first map DMP questions to their corresponding FAIR principles. We then filter out which questions from the mapping can be answered using suggestions from FIPs. For RQ2, we create a DMP interface on the FAIR Wizard platform[5] with text-based suggestions. Finally, for RQ3, a user study is conducted followed by a survey to understand how participants take suggestions from FIPs. This paper made the following research contributions.[6]

1. A generic workflow for using FIPs in the interface as suggested in DMPs.
2. A mapping between the chosen DMP template and the FAIR principles.
3. An analysis of the relationship between the DMP template and the FIPs.
4. A reusable and extendable knowledge model in the FAIR Wizard platform.
5. A user study aimed to understand how researchers can effectively use FIPs as suggestions while creating DMPs.

The paper is organized as follows: Section 2 provides an overview of related work. Section 3 gives an overview of the workflow. Section 4 provides more details about DMPs and FIPs and identifies which questions in the chosen DMP template can be addressed using FIP suggestions. Section 5 delves into the creation of an interface on the FAIR Wizard platform. Section 6 provides details of the user study. The evaluation of the resulting DMPs and the survey is presented in Section 7 followed by some discussion in Section 8. Finally, Section 9 presents the conclusion and outlines future work.

2   Related Work

Despite the potential impact of FAIR community standards on the researchers' choices over management and publication of research data, the connection between DMPs and FAIR principles has been examined empirically only in a few studies. Henning et al. [2] analysed 10 DMP templates and concluded that DMPs fail at capturing detailed community-specific implementations, especially of the Interoperability principle, and do not cover metadata management sufficiently. In a study based in the US [5], DMPs associated with grant proposals were analyzed alongside interviews to the PIs who wrote them. The author found that the more technical parts of the DMPs (including questions on FAIR) were the least detailed, and that PIs would need more training and guidance on the more specialised concepts such as FAIR and metadata. Likewise, a report by OpenAIRE on the Horizon 2020 Template For Data Management Plans highlighted the need to clarify issues and terms around FAIR implementation [1]. Finally, in a paper outlining how to use the Open Science Framework (OSF) platform for

---

[5] The FAIR Wizard uses the Jinja template engine: `https://fair-wizard.com/`.

[6] The DMP template, the knowledge model, the mock DMPs, the data of the survey, the analytical results, and a video demo are on Zenodo: `10.5281/zenodo.10286752`. The Python scripts are at `https://github.com/FAIR-Expertise-Hub/FIP2DMP`.

DMPs, Sullivan et al. [10] indicated the importance of referring to best practices applied to different research contexts. All in all, studies seem to agree that having clearer examples about FAIR implementation and standards used in other projects would be beneficial for researchers filling in DMPs, however this has not been tested systematically.

To our knowledge, the only attempt explicitly link the DMPs with the FAIR community standards declared in FIPs used the DMP template of the Leiden University. The authors identified seven questions in their DMP template that could be linked to the FIP questions [3].[7] They proposed to develop a knowledge model (KM) of the Leiden University DMP template and import answers from some FIP as pre-filled answers, leaving it to the user to select the most relevant ones from relevant communities [7]. A KM is a structured document with questions, explicative text accompanying each question, and the type of answers specified. The template includes examples of good data management practices and guidance on how to meet the requirements of different funding agencies and institutions. Our research is inspired by this work and follows the same approach for the handling of DMP template on FAIR Wizard but differs in the DMP template used and how (multiple) FIPs are handled as suggestions. Moreover, we include a user study.

3   Workflow
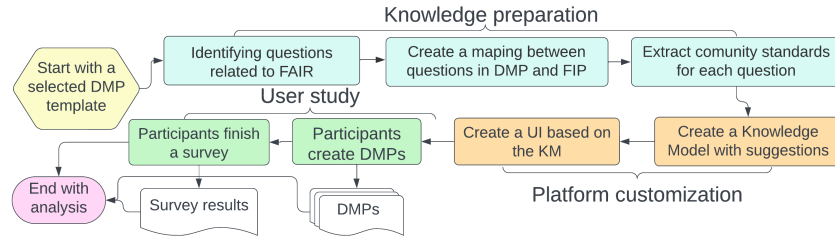
Fig. 1. A visualization of the workflow of our approach with different colors for the three stages: knowledge preparation, platform customization, and user study.

Given a DMP template, to provide suggestions for the right questions in DMP, we need to find the mapping between the questions in DMP and FIPs. We use the FAIR principle as an intermediate step to narrow down this correspondence. Once obtained, the mapping is used to associate with selected questions in the DMP some extracted community standards captured by the FIPs. A KM is then constructed to be loaded to the FAIR Wizard for a customized interface. Each study participant is asked to create a DMP and complete a survey. Figure 1 is a visual representation of this workflow.

---

[7] The mapping is at `https://osf.io/5jsfp`.

4   Knowledge preparation: Connecting DMPs with FIPs

Next, we explain the details of DMPs (Section 4.1) and assign the FAIR principles to the questions in DMPs (Section 4.2). Following that, we introduce the FIPs in detail (Section 4.3) and presents the mapping between DMP questions and FIP questions (Section 4.4).

4.1   Data Management Plans

DMPs have become a standard in recent years and are required more and more by funding organizations. Miksa et al. describe them as 'awareness tools', shedding light on data management practices employed in research projects [6]. In practice, they assist researchers in ensuring proper management, documentation, and preservation of data, while also meeting funders' requirements. DMPs are implemented from DMP templates, which consist of questions with specified answer types. DMP templates can be very different from each other, as they are often tailored to best facilitate the data management of researchers and meet the requirements of the funding organizations. Various tools can be used to instantiate such templates with a user interface, such as DMPOnline[8], DMP Tool and Data Stewardship Wizard [9]. For this paper, we use the DMP template by the Vrije Universiteit Amsterdam (VU). The template is hosted on DMPOnline. Researchers can easily fill in their answers and download the completed DMP in the PDF and JSON format.

DMPonline's strength lies in the relative ease for researchers to start filling in their data management plan and request feedback. The data stewards, on the other hand, can easily provide comments and guiding article on the side line. This is a fairly manual interaction, though, in which hardly any automation is present. DMPonline automatically provides the data management plan templates of all major funders of Europe, including NWO, ERC and ZonMW. These templates differ in accessibility for beginning researchers, with most of the templates focusing specifically on the FAIR needs of the funder. The NWO template, for example, covers 18 questions with multiple sub-questions bundled together. It directly addresses FAIR principles, asking researchers how they will ensure data FAIRness. The template provided by ERC consists of 8 questions, emphasizing FAIRness by asking researchers to describe how they will make their data findable, accessible, interoperable, and reusable. However, such broad questions make it difficult to associate with a specific FAIR principle, which may hinder machine-actionability. Therefore, they are not to be considered for this research. Years of iteration and refinement have made the current VU DMP template to cover all the requirements by major funders, as well as guide the researcher in planning the research data management required to streamline their research at the VU. The result is a template with roughly 53 questions, depending whether the research will contain personal or sensitive data. Guiding instructions are provided along the step-by-step questions. The current DMP template covers

---

[8] https://DMPOnline.vu.nl/

the two formerly parallel editions which are made for research with and without personal data, incorporating questions omitted based on researchers' choices in the document along with guiding text. Furthermore, DMPOnline dynamically displays or omits questions based on the researcher's responses regarding dealing with personal data. By March 2023 when this project started, the DMP in use at VU is identified as '*1 - VU DMP template 2021 (NWO & ZonMW certified) v1.3*'. This version consists of 53 questions[9] and is developed and maintained by the data stewards of the University Library. The VU DMP template is used by all faculties, with exception of the Medical Faculty. The Medical Faculty maintains their own template and employs a separate data management team for their research support. In this study, we focus on researchers in the Faculty of Social Sciences. The questions in VU's DMP template cover a wide range of topics including authorship, legal and privacy ethics, funding number, etc.

When comparing the VU DMP template with the current Leiden version,[10] there are notable differences. The Leiden DMP consists of 48 questions (compared to 53 questions of the VU template), mostly in multiple-choice format, in contrast to that of the VU. Leiden places more emphasis on privacy concerns and security risks.

4.2   Mapping the DMP Template to the FAIR principles

Since the FAIR principles are widely endorsed for good data management, most DMPs also include questions on the implementation of FAIR [2]. Hence, to answer RQ1, we begin by analyzing the 53 questions in the chosen VU DMP template and exclude questions that do not correspond to any FAIR principle. After consulting the team at the university library that develops and maintains the DMP and compare it against the FAIR principles, we identified a total of 17 questions relevant to the FAIR principles.[11] More specifically, 14 of the questions are about the Findability principle, and two questions are related to the Accessibility principle. Three other questions focus on the Reuse principle. Since the template was designed without exactly following the FAIR principles, questions can correspond to multiple principles. For example, Question 5.1 corresponds to both F2 and R1.2 principles. Among all these questions have to do with the FAIR principles, we observe that 82.4% of the questions are about Findability since a significant amount of questions are to do with persistent identifiers. A comparison with the past attempt is in Section 8.

4.3   FAIR Implementation Profiles

A FAIR Implementation Profile (FIP) is a set of choices made by a FAIR Implementation Community (FIC) on how to implement the FAIR principles [8]. The

---

[9] It was the latest version at the time of project, consisting of the most amount of questions following the division of metadata questions into questions that follow a more manageable line of thinking.

[10] The Leiden DMP template used is at `https://zenodo.org/records/4423065`.

[11] Details are given in the supplementary material.

community decision in a FIP is collected using a questionnaire[12] in which experts and members of the community collectively indicate their preferred FAIR Enabling Resources (FERs), FERs are the technologies and standards that are FAIR by themselves, that are needed to achieve a specific aspect of FAIR implementation [8,11]. Each question corresponds to a FAIR principle. Decisions made by communities can significantly impact researchers, influencing choices related to data repositories, licenses, and metadata standards. FIPs have been conceptualized to foster convergence of FAIR implementation efforts across communities and domains [8]. FIPs are filled in on the FIP Wizard[13], an online platform developed by the Data Stewardship Wizard together with the GO FAIR Foundation[14]. The FIP Wizard provides an easy-to-use interface and allows to publish the FIPs in a machine-actionable format as nanopublications.

There are already over a hundred FAIR Implementation Communities (FICs) covering various domains. For this study, we focus on the only six FIPs that pertain to communities in the social sciences [11]. The social sciences constitute an interesting use case due to their longstanding tradition of data sharing, showed for instance by the availability for research purposes of large-scale survey data, combined with a large heterogeneity in the standards adopted. Among the six FIPs, three stem from communities that publish survey data: GESIS social Science Survey Research (GESIS SSSR), the European Social Survey (ESS), and the Australian Correspondent (AUSSI-ESS). In addition, the Dutch Socio-Economic History (SEH) and the Media Content Analysis Lab (MCAL) are two communities in the Netherlands. Lastly, the LGBTQ+ Linked Open Vocabulary (LGBTQVoC) community creates multilingual LGBTQ+ controlled vocabularies for indexing digital records to represent LGBTQ+ objects in non-English languages. Due to the page limit, details about these communities, the creation of the FIPs and their details, and how they can be used for analysis of convergence and comparative studies are omitted but can be found in [11].

Table 1 compares the FIPs on the FAIR principles that are also relevant for DMPs. The row for 'F1 Data' highlights convergence among social science communities, indicating a preference for DOI. Regarding 'R1.1 Data', three out of six communities indicate the use of different versions of CC-BY licenses for their data. This information could indicate the popularity of certain resources across/within selected communities.

However, the examination of FERs regarding other principles shows the differences. Many different FERs are mentione under 'F2': for instance, BIBFRAME is highly specific to LGBTQVoC, whereas DDI-Codebook is mostly mentioned by communities that deal to some degree with survey data. The FERs indicated under 'F4 Data' show fragmentation, with each FIC indicating different data repositories. Since these FIPs are all distinct despite being in social science domain, it is important that the researchers carefully consider which community aligns the most to their research project's objectives to maximize findability and

---

[12] https://bit.ly/yourFIP

[13] The FIP Wizard platform: https://fip-wizard.ds-wizard.org

[14] https://ror.org/056j50v04

FAIR Principle	FIP question	SEH	MCAL	LGBTQVoC	SSSR	ESS	AUSSI-ESS
F1 Data	What globally unique, persistent, resolvable identifiers do you use for datasets?	DOI (future), Handle	DOI	DOI	-	DOI, DDI URN	DataCite DOI resolution service
F2	Which metadata schemas do you use for findability?	MARC21, EAD3, DDI-Codebook, DCAT2	-	MARC21, BIBFRAME	DDI-Codebook	DDI-Lifecycle 3.3	DDI-Codebook, DataCite 3.1
F4 Data	In which search engines are your datasets indexed?	Dataverse, DANS SSH data station (future)	DANS SSH data station, Dataverse, OSF, Figshare, TriplyDB (future)	-	GESIS Search	ESS Data Portal, EOSC Portal	ADA Dataverse FER
R1.1 Data	Which usage license do you use for your datasets?	CC-BY-SA, CC-BY-NC	-	CC-BY-NC-ND 4.0	GESIS Usage Regulations 2018	CC-BY-NC-SA 4.0	-
R1.2 Data	Which metadata schemas do you use for describing the provenance of your datasets?	DCAT2	-	-	-	DDI-Lifecycle 3.3, DDI-Codebook, DDI-CDI (future)	DDI-Codebook

Table 1. A comparison of decisions on the practice of the FAIR principle

reusability of their research data by other researchers in the same community. Finally, the blank answers could make the suggestion less informative.

4.4   Assigning FIPs to DMP Questions

Next, we use the mapping in Section 4.2 and find the correspondence between questions in FIPs and DMPs. We manually examined questions that correspond to the same FAIR principle and identified eight DMP questions that may be successfully connected to FIPs. However, one question (Question 6.3) was ex-

ID	DMP Question	FAIR principle	FIP Question
4.6	Where will you publish your data assets?	F4 Data	In which search engines are your datasets indexed?
4.8	How will you ensure your data assets get a persistent identifier (e.g. a DOI-code)?	F1 Data	What globally unique, persistent, resolvable identifiers do you use for datasets?
4.9	Will you register your datasets in an online registry other than PURE? If yes, where?	F1 Data	What globally unique, persistent, resolvable identifiers do you use for datasets?
4.13	Please indicate the license and/ or terms of use under which you share your data.	R1.1 Data	Which usage license do you use for your datasets?
5.1	What metadata and documentation will accompany the project?	F2 & R1.2 Data	Which metadata schemas do you use for findability? & Which metadata schemas do you use for describing the provenance of your datasets?
5.2	What metadata and documentation will accompany the data assets?	F2	Which metadata schemas do you use for findability?
5.3	What methods, software or hardware are needed to access and use your data?	R1.2 Data	Which metadata schemas do you use for describing the provenance of your datasets?

Table 2. Mapping of DMP Questions to FIP questions via the FAIR principles.

cluded due to its ambiguity.[15] Table 2 illustrates the correspondence between these DMP questions to the FIP questions after manual examination. Question 4.6, 4.8, 4.9, and 4.13 have a clear one-to-one mapping to the corresponding FIP questions. The case of the next section is more complicated. Question 5.2 focuses on metadata schemas for data assets. Question 5.3 addresses provenance models and methods. Question 5.1 bridges both, encompassing rich metadata and detailed provenance for the entire project, with metadata schemas. The next section shows how answers in the FIPs are included in the customized interface.

5   Building an interface on the FAIR Wizard

Fig. 2. FERs and corresponding communities based on user-selected FIPs in FAIR Wizard

Fig. 3. User describing answer in text and selecting FER from a list of FERs using search engine in FAIR Wizard

To address RQ2, in this section, we explain how the KM is constructed and the corresponding interface on the FAIR Wizard platform. Despite that VU's current DMP is hosted on the DMPOnline platform, no suggestion can be provided and the platform does not support customization of its interface. Furthermore, the platform cannot convert the resulting DMP to other machine-actionable formats and does not support loading content from other datasets using customized queries. Thus, we migrate to the FAIR Wizard, which covers

---

[15] Question 6.3 'For data that are only available upon request, what methods will be used to handle requests for access and how will data be made available to those requesting access?' was linked to FAIR principle A1.2 in the previous step.

these weakness. Developers can create a KM that specifies the structure of the DMP, the type of answers, and some customized functions to retrieve external information. Suggestions extracted from chosen FIPs are included as additional information below the questions. For each question, two types of information are expected: a list of FERs and a string description that explain how the chosen resources will be used and some additional information including resources not found. In this paper, we limit our KM to only questions chosen in Section 4.4. The created KMs can be modified, downloaded, and reused. Finally, the interface is automatically generated when the KM is uploaded.

Figure 2 illustrates that below the question and its description, a two-column table that specifies the FAIR Implementation Community (FIC) and the FAIR Enabling Resources (FER) is provided as suggestion. The left column of the table presents all the decisions made by research communities, while the right column shows the corresponding research communities that made the decisions. For instance, the table shows that SEH uses the Dataverse Project to publish their data assets. When the user takes a suggestion from a research community, they can then search for the corresponding FER as shown in Figure 3. In this case, when typing in 'dataverse', various related resources are displayed.[16]

6   User Study

After reaching out to almost all the researchers in the Faculty of Social Sciences of the VU[17], a total of six researchers agreed to participated in this study. These participants have either written some DMPs or have sufficient knowledge about DMPs. Participants first pick one DMP that is closest to their research among four mock DMPs. These mock DMPs were created based on some ideas of existing DMPs by university researchers in the past[18]. They were then asked to complete the DMP by using the interface as described in Section 5. The user study ends with a survey.

The survey comprises of twelve questions in two parts. In Part A, participants assess the relevance and helpfulness of the suggestions. More specifically, this section covers evaluating community relevance for DMPs, assessing decision

---

[16] A video demo of the interface is included in the supplementary material.

[17] Teaching staff, supporting staff, retired professors, external and visiting researchers were excluded. PhD students were included.

[18] For this user study, we newly developed four mock DMPs based on relevant existing DMPs, covering various research topics and facilitate potentials for the use of FERs and simulate real-world scenarios in research data management. To focus on the aspects under evaluation, irrelevant text was omitted in these mock DMPs. Participants could select a mock DMP on a research topic they are familiar with. Questions 4.6 and 4.8 have some pre-filled answers based on our understanding of the research topics to provide participants a starting point. That of 4.9, 5.1, 5.2, and 5.3 are not provided with suggestions but Question 4.13 has been pre-filled using the VU's default licensing option, which is CC-BY 4.0. The participants shall make their own independent decisions in response to these questions.

alignment with communities, and evaluating the effectiveness of suggestions for specific DMP questions, along with the ease of locating FERs. In addition, Part B focuses on the background and experience of participants. They need to indicate how long they are in their academic career and how much experience they have with DMPs. Additionally, they are asked to rate the clarity of the study's goals. Finally, the survey includes a question about how they consider aligning with community standards in comparison with other stakeholders (e.g. university requirements, journals). The five most relevant questions are listed in Table 3. The remaining questions and other details are given in the supplementary material. The user study was completed in October 2023.

7   Evaluation

To address RQ3, in this section, we summarize the users' backgrounds in Section 7.1, the efficacy and significance of research communities in Section 7.2, and associating FERs with DMPs in Section 7.3.

7.1   Participants' Background and Experience

For our analysis, we used the answers of five participants[19]. The participants had a range of 3 to 10 years of experience in academia, counting from the start of their PhD. With the exception of one participant, all participants had prior experience in writing DMPs. The survey results indicated that participants, on average, found the objective of the study to be moderately clear, with a mean rating of 3.4 out of 5 (see Q10 in Table 3). Some indicated the lacking of knowledge about FIP and FER.

7.2   The Efficacy of Suggesting Community Standards in DMPs

Due to page limit, only some key survey questions and the corresponding results are included in Table 3, covering aspects like community relevance, DMP suggestion helpfulness, alignment consideration, study goal clarity, resource search ease, academic experience, previous DMP writing, entity impact, and challenges faced during DMP creation. In Question 1, participants were asked to rate the relevance of research communities for their DMPs. The mean rating of 3.4 out of 5 suggests that, on average, participants found research communities to be relevant to their DMPs. Responses of Q2 spanned from 2 to 4, with a mean rating of 3.4 and a median of 4, indicating that, participants perceived these suggestions as helpful. As for Q3, participants found alignment with community decisions to be of moderate importance.

Next, we evaluate the effectiveness of suggestions for each DMP question. Figure 4 shows that questions 4.6, 4.8, 4.9, and 4.13 in the VU DMP template were

---

[19] The PDF file of one of the mock DMPs downloaded from the DMPOnline had missing pages. This was not reported until the end of the user study. One participant chose that DMP and was therefore excluded from the analysis.

Survey Question ID	Survey question	Mean	Median	Std
Q1	On a scale of 1 to 5, how relevant are communities for this DMP? 1 indicating that no community is relevant and 5 indicating that many communities are relevant.	3.4	3	0.89
Q2	On a scale of 1 to 5, please evaluate whether the suggestions provided in this DMP are helpful for the communities in answering their corresponding questions. 1 indicating that it is not helpful and 5 indicating that it is very helpful.	3.4	4	0.89
Q3	On a scale of 1 to 5, how much would you consider aligning the decisions in this DMP with those made by the relevant community? 1 indicating minimal alignment and 5 indicating complete alignment.	3.4	3	1.14
Q7	On a scale of 1 to 5, how easy was it for you to find the FAIR-Enabling Resource in the search bar? 1 indicating extremely difficult and 5 indicating extremely easy.	3.6	4	1.14
Q10	On a scale of 1 to 5, how clear was the goal of the study to you? 1 indicating not clear at all and 5 indicating very clear.	3.4	4	0.89

Table 3. Survey questions and their results together the range, mean, median, and standard deviation.

generally perceived as helpful or inspiring to some extent. However, a closer look reveals varying responses to question 5.1. This question, which inquired about the types of documentation to be produced during the research project, was frequently perceived as not helpful or misleading. This might be attributed to the inherent ambiguity of the question, but also the weak link with the FIP question, which does not include the documentation part. The DMP question demands descriptions of documentation, including codebooks, lab journals, read-me files, research logs, and protocols. The challenges arise because the FIP question can only provide FERs. Questions 5.2 and 5.3 were less attended with neutral results.

In addition, our analysis shows that many FERs from SEH and MCAL communities were selected. On the other hand, few FERs that are specific to the SSSR community were chosen. Due to the limited number of participants and the diversity in mock DMPs, we cannot conclude the most popular community that participants attempt to align to.
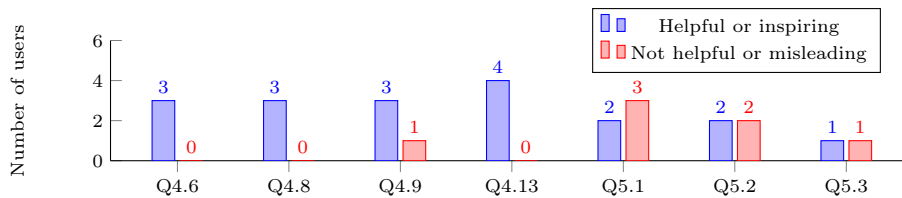
Fig. 4. Comparing the impact of suggestions on DMP questions: helpful or inspiring vs. not helpful or misleading.

7.3   Associating FAIR Enabling Resources with DMPs

In this section, we focus on participants' experience with FERs. Only one participant reported successfully finding all the FERs they intended to specify. The other reported missing FERs. However, when asked to rate how easy it was to find FERs using the search bar, the mean rating is 3.6 out of 5. This suggests that although most participants cannot find all the FERs they wanted through the search bar, they find the search bar easy to use for finding resources. Moreover, An analysis of the resulting DMPs shows that all participants managed to specify at least one FER for questions 4.6 and 4.13. Additionally, questions 4.8, 5.2, and 5.3 were frequently specified through the search bar. Conversely, for questions 4.9 and 5.1, most users struggled to locate the desired FERs through the search bar. This is not surprising, as the participants find the suggestions for Question 5.1 not helpful or misleading (as described in Section 7.2).

8   Discussion

It can be ambiguous when aligning questions in the VU DMP template with the corresponding FIP questions via the FAIR principles. This is because the chosen template did not take the FAIR principles into account by design. As a result, the suggestion provided can lead to confusion. Our analysis shows that, out of the seven identified DMP questions, three questions (5.1, 5.2, and 5.3) face the problem of providing unhelpful or misleading suggestions. This is because the DMP questions 5.1 and 5.3 cover multiple issues and are only partially linked to the FIP questions, requiring additional explanation. This calls for a revision of some questions, especially 5.1 and 5.3. Moreover, both correspond to R1.2 where three FIPs lack resources to contribute to the suggestion. In addition, Question 4.9 assumes that researchers will be registering their datasets on the PURE system (required by the university). However, the formation of the question excludes PURE as an answer, which could lead to a missing FER and can be a problem for automated analysis at a later stage. Moreover, at further inspection, it emerged that principle F4–Data would match this DMP question better, yet this was missed due to the way the FIP question was formulated, mentioning search engines. It is only recently that the FIP question was re-formulated and now focuses on services (including registries). This shows that an accurate mapping is crucial for FIP suggestions in DMPs to make sense. Moreover, the way FIPs interpret 'Accessibility' focuses mostly on machine accessibility. In contrast, DMPs focus more on human accessibility for repositories to make machine accessibility possible. This can lead to confusion and result in some missing pairs in the mapping.

Despite the detailed introduction with used terminology explained on the first page of the interface, some participants expressed that they found it difficult to comprehend certain terminologies and lack of knowledge about metadata standards and DMPs. This could also be a cause of confusion for Question 5.2. In practice, metadata handling is typically the responsibility of data repositories or data stewards. In this study, participants cannot consult data stewards.

As a proof-of-concept work, six FIPs in social sciences were used. However, there are missing FERs for some questions. In such cases, participants would face the situation with missing or incomplete guidance from these communities and would have to independently formulate their responses. If used in practice, this may lead to the adoption of standards from an alternative community, which could potentially diminish the findability and reusability of their research data.

Furthermore, the current KM on the FAIR Wizard platform generates a fixed project template and lacks the functionality to customize the selection of FIPs. Thus, users are presented with suggestions uniformly based on all available FIPs. This process could be automated. If users can choose their relevant research community beforehand, the suggestions could be clearer and more domain-specific.

Inspired by [4], we included in the survey a question about which stakeholders have the most impact on their decision-making. Our participants indicated that the department, faculty, and the university research data management team, as well as the ethics committee, have the most significant influence on their decisions in DMPs (13.95%). That of community shares the second place with the university I.T. team, as well as the data management platform (9.30%). Some other factors would also be taken into account. Despite the scale of the survey, it shows that the decision can be influenced by many factors. Community standards are important, but maybe not among the most influential factors. However, these suggestions and guidelines are mostly formatted as textual information in PDF format, thus cannot be easily imported into a DMP editing system.

Finally, when comparing our study to the research using the Leiden University's DMP template [3], we encountered the same challenge in addressing interoperability. Despite both studies identifying questions in their respective DMP templates that can be linked to FIP questions, we observed that our study, based on the VU DMP template, included a slightly larger subset of questions (7 compared to Leiden's 6). They proposed to use a "project importer" feature to pre-fill answers in the DMP using existing FIPs, which is different from our approach. Given the above-mentioned factors and multiple FIPs relevant for a DMP, pre-filling an answer could cause confusion. Both studies address several issues with respect to the DMP templates used.

9   Conclusion and Future Work

In conclusion, this paper explored how FIPs can be used as suggestions for DMPs and whether researchers can align their DMPs with community decisions through the use of FIPs. To address RQ1, we constructed a mapping between DMP and FIP, and identified seven DMP questions that could be effectively addressed using community standards captured by FIPs. For RQ2, we constructed a KM tailored to the VU DMP template with the information of six distinct research communities' standards integrated into each question. Finally, for RQ3, a user study was conducted, which revealed that, for some questions, users find suggestions from FIPs helpful or inspiring. This is not the case for others. Therefore, some discussion was provided.

In future research, we plan to enhance our KM with customization features, allowing users to pre-select the FIPs from which they wish to follow the decisions. As a proof-of-concept, our primary focus is on the social science using six FIPs,but we aspire to broaden the scope of our work to include other domains and FIPs in the future. In future studies, mock DMPs could be replaced by researchers' own (past) DMPs.

Furthermore, researchers' answers are often embedded in text in most existing DMP tools. Essentially, the introduction of FER to DMP is a step towards bringing DMPs to the ecosystem of linked data. Although the RDA DMP Common Standard has taken a significant step towards creating machine-actionable Data Management Plans (maDMPs),[20] there is still a gap until DMPs become FAIR objects in the semantic web. When fully integrated, the DMPs could then take suggestions from not only FIPs but also other types of resources, which in turn, can be used to inform the creation of new FIPs.

Finally, a detailed review of the DMPs is essential for future research connecting DMPs with other resources. Catching up with other DMPs [2], especially for what concerns metadata, 'Accessibility', and 'Interoperability', is needed.

References

1. Marjan Grootveld et al. OpenAIRE and FAIR Data Expert Group survey about Horizon 2020 template for Data Management Plans. 2018. 10.5281/zenodo.1120245.
2. Patrícia Henning et al. The FAIRness of data management plans: an assessment of some european dmps. *Revista Eletrônica de Comunicação, Informação e Inovação em Saúde*, 2021.
3. Kristina Hettne et al. Fip2dmp: Linking data management plans with fair implementation profiles. *FAIR Connect*, 1:23–27, 01 2023.
4. L. Kvale et al. Understanding the data management plan as a boundary object through a multi-stakeholder perspective. *Int. J. Digit. Curation*, 16:16, 2021.
5. Sara Mannheimer. Toward a Better Data Management Plan: The Impact of DMPs on Grant Funded Research Practices. *Journal of eScience Librarianship*, 7(3), 2018.
6. Tomasz Miksa et al. Ten principles for machine-actionable data management plans. *PLoS Comput. Biol.*, 15(3):e1006750, 2019.
7. Fieke Schoots et al. Data management plan template leiden university (4.2). Zenodo, 2021. `https://doi.org/10.5281/zenodo.4423065`. Accessed June 2023.
8. Erik Schultes et al. Reusable FAIR Implementation Profiles as accelerators of FAIR Convergence. In Georg Grossmann and Sudha Ram, editors, *Advances in Conceptual Modeling*, pages 138–147, Cham, 2020. Springer International Publishing.
9. Jurrian H. Spaaks et al. The data stewardship wizard: A tool bringing together researchers, data stewards, and data experts. *Int. J. Digital Curation*, 14(1):165–175, 2019.
10. Ian Sullivan et al. Open and reproducible research on open science framework. *Current Protocols Essential Laboratory Techniques*, 18, 2019.
11. Shuai Wang et al. FAIR Implementation Profiles for Social Science. In *Proceeding of 17th Int Conf on Metadata and Semantics Research*. Springer, 2023.

---

[20] `https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard`