



OpenAIRE and the FAIR Data Expert Group

Report on the results of the survey about the
Horizon 2020 template for Data Management Plans



Authors

Sarah Jones¹, Ellen Leenarts², Marjan Grootveld³, Eliane Fankhauser⁴, and Emilie Hermans⁵

Note

The final report, together with a summary infographic, the full survey responses and raw/analysed data are published in a collection on Zenodo under a public license with the following DOI: 10.5281/zenodo.1120245.

Rights



Creative Commons Attribution 4.0 International License (CCBY 4.0)

This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

¹ DCC, University of Glasgow, G12 8QQ Glasgow, United Kingdom, sarah.jones@glasgow.ac.uk, <https://orcid.org/0000-0002-5094-7126>

² DANS, 2593 HW Den Haag, The Netherlands, ellen.leenarts@dans.knaw.nl

³ DANS, 2593 HW Den Haag, The Netherlands, marjan.grootveld@dans.knaw.nl, <https://orcid.org/0000-0002-2789-322X>

⁴ DANS, 2593 HW Den Haag, The Netherlands, eliane.fankhauser@dans.knaw.nl

⁵ Ghent University Library, University of Ghent, 9000 Ghent, Belgium, emilie.hermans@ugent.be



Table of Contents

Authors.....	2
Note	2
Rights	2
Introduction to the survey.....	5
Summary and recommendations	6
General outcomes.....	6
Improvements suggested by respondents and developments	7
Recommendations: Restructuring the DMP template	9
Recommendations: Tailored guidance, examples & terminology.....	9
Recommendations: Support on costing and rewards	10
Recommendations: Clarifying the review process	10
Outcomes per question	11
Question 1 - Role of participants	11
Question 2 - Name and affiliation.....	12
Question 3 - Reaction to H2020 DMP process.....	12
Question 4 - Experience with H2020 DMP template	13
Question 5 - Unknown, unclear or confusing terminology.....	14
Question 6 - Confusing, inconsistent or redundant questions	16
Question 7 - Missing or potential guidance.....	17
Question 8 - Topics that are missing or should be removed	19
Question 9 - Issues encountered following EC guidelines	20
Question 10 - Suggestions on the DMP template.....	21
Question 11 - Priorities for a DMP template or tool	23
Question 12 - Openly publishing DMPs	25
Question 13 - Resources to include in disciplinary guidance.....	26
Question 14 - Suggestions to EC on DMP approach	27
Question 15 and Question 16 - Providing feedback on DMPs.....	29
Question 17 - How H2020 DMP support process differs.....	31
Question 18 - Time needed to review H2020 DMP	32
Question 19 - Other recommendations on FAIR data management.....	33
Appendix I - Infographic of the survey results	35
Appendix II - Recommendations and next steps	36
1. Clarify EC requirements for DMPs	36



2. Revise the DMP template structure.....	36
3. Simplify the DMP content and terminology.....	37
4. Provide discipline-specific guidelines and example answers.....	37
5. Encourage the publishing of DMPs and collate examples.....	37
6. Facilitate the inclusion of RDM costs in grant applications	38
7. Improve DMP review practices and share guidelines	38
Appendix III - List of survey questions	40
Appendix IV - Request to circulate the survey.....	43



Introduction to the survey

The OpenAIRE project supports the open science ambitions of the European Commission. The project and in particular the Research Data Management team provide support, training and information on the [Open Research Data Pilot](#). In this context, a survey was carried out to collect feedback on the Horizon 2020 template for Data Management Plans (DMPs). The team collaborated with the [Fair Data Expert Group](#), which is providing recommendations to the European Commission on turning FAIR data into reality. One of the specific tasks of the Expert Group is contributing to an evaluation of the Horizon 2020 approach to DMPs, including future revisions of the template and the development of additional sector/ discipline-specific guidance. The aim of the survey was to collect experiences of researchers and DMP reviewers with the DMP template and guidelines on FAIR data management in Horizon 2020. The survey assesses the usefulness of the guidelines and any aspects that are confusing and unclear to determine what improvements can be made.

Feedback was sought from both researchers and research support staff. The survey was initially scheduled to run from 22 May to 21 June 2017. Several organisations were asked to help announce the survey, including OpenAIRE's [National Open Access Desks](#), the Fair Data Expert Group, [FOSTER](#), [LIBER](#), and the [RDA Interest Group on Active DMPs](#) (see Appendix II). When the first survey responses showed only a small share of researchers, more stakeholders were contacted to specifically target this community. The [European Research Area](#) was approached, whose project officers circulated the survey call among award holders of EC projects. Early-career researchers were also informed through the [YEAR network](#) and [EURODOC](#). This resulted in an extension of the survey to 21 July 2017.

At the close of the survey on 21 July 2017, a total number of 289 responses were reached. 50% of the respondents indicated that they were researchers, and 60% that they were (also) research support staff. OpenAIRE and the Fair Data Expert Group are very pleased with this balanced outcome and would like to thank all colleagues and organisations who promoted the survey, as well as everyone who took part in it.

The next section summarises the outcomes and provides recommendations based on the responses. After this the report addresses responses per question. This report concludes with appendices containing an infographic that summarises key findings, the email announcement used to circulate the survey, and a list of all survey questions.

Together with this report, a summary infographic, the full survey responses and raw/analysed data are published in a collection on Zenodo with following DOI: 10.5281/zenodo.1120245 and link: zenodo.org/record/1120245.



Summary and recommendations

In summer 2017, OpenAIRE and the Fair Data Expert Group ran a survey on the European Commission's approach to Data Management Plans (DMPs). Nearly three hundred people responded, of which 50% identified themselves as researchers, project coordinators or others who write DMPs for the European Commission [Question 1]. Respondents overwhelmingly found the process of writing a Horizon 2020 DMP positive (60% positive, 16% negative, 24% not applicable [Q3]), despite initial reservations in some cases. Several comments indicated that while it seemed an administrative burden at first, the process of developing the DMP opened up useful discussions and caused projects to reflect on potential issues that would have been overlooked otherwise. A small cohort also concluded that it was unnecessary and bureaucratic, indicating that clearer guidelines are perhaps needed as to when DMPs are not required (e.g. fellowship programmes or when no data is created).

General outcomes

The concept of FAIR was felt to be well-understood (44% strongly agree [Q4]), and the majority thought the EC's DMP template was very useful (45% agree [Q4]). That said, the free-text questions [Q5-Q10] probing on terminology issues, redundant questions and potential improvements, elicited many detailed responses, indicating that while the concepts work at a high-level, many questions remain around implementation. Interoperability is the main term that caused confusion, closely followed by metadata, standard vocabularies and ontologies. More remarks on FAIR were made in the context of Q19 eliciting general recommendations. Two other terms that caused issues for several respondents were data access committees and certified repositories [Q5].

There were a number of references to local RDM support teams, national data services and project officers who clearly helped to interpret the requirements and apply them to the researchers' context. Some commented that without this kind of help, they would not have been able to complete the DMP. Ensuring that each of these groups has access to the latest materials is critical, as several comments were made about misinformation. Further internal training courses should also be run to help project officers advise beneficiaries and review DMPs. Using networks such as the OpenAIRE NOADs and RDA-Europe will be key to disseminate up-to-date information on the EC's requirements to the wider RDM support community.

Preservation and legal issues were felt to be underplayed in the template, while several questioned the inclusion of ethical issues since there are separate requirements on this. The template was felt to be too structured and prescriptive to allow for different types of projects, and a few people commented that using each section heading was sufficient in most cases. The use of FAIR to structure the template posed several issues as concepts overlap and questions ask for similar information in an unhelpful order which makes it difficult to avoid repetition. For example users are asked to explain how they will make their data findable before discussing the chosen repository.



The timing of the DMP applied by the EC was generally felt to work well, as it is only actual projects that are required to develop a plan. However, practicalities around updating DMPs and including RDM costs need some thought. Several projects referenced issues with the DMP being a living document – it was unclear how often it should be updated and what level of information was expected at each stage, some project officers were unaware updates were needed, and Commission systems often meant deliverables had to be rejected to allow updated versions to be submitted. More clarity was also called for on data management costs that are eligible for reimbursement and how to incorporate these into proposals as DMPs are not required until month 6 of the project. Asking a few key questions at the application stage to prompt inclusion of costs would be worthwhile.

Improvements suggested by respondents and developments

One of the survey questions proposed several potential improvements to help guide further work in this area. We highlight this question and the respondents' preferences here because they relate to several key topics, as well as to current developments. More information on the figure below can be found in the section discussing Question 11.



Figure 001 - Q11: priorities for a DMP template or tool

The top priority improvement to DMP templates or tools was to suggest relevant standards for the field and data type, receiving 101 votes out of 652 votes in total (each respondent could choose up to 5 improvements - the number behind each bar shows a weighted average, not the number of votes). Many votes were also cast for providing example answers, providing dropdown options and recommending repositories. User needs are at present focused on obtaining more tailored advice and support to ease the process of developing a DMP. Several features that require data exchange across systems such as alerting repositories or data service providers to support requirements and pre-filling information in the DMP were also proposed. Of these, pre-filling or automatically updating information was the most called for (63 votes), followed by sharing information with university or data services (45 votes).

Work has been started by Science Europe [to develop domain protocols](#) that provide guidelines and predefined responses for different fields. There is potential to overlay these **discipline-specific guidelines** onto the EC's common template, providing different options for researchers to select from. The DCC has also been integrating the [RDA Metadata Data Standards Directory](#) and [FAIRsharing](#) databases into [DMPonline](#), again providing more structured options to direct user responses. For the long-term preservation of and access to data we recommend researchers to use a domain repository when available; if not, generic repositories like Zenodo and Github are worth considering.

A high-priority request for improvement is to **provide example DMPs**. Ideally, a set of actual, approved H2020 DMPs would be published that covered a wide range of disciplines and project types. This would allow researchers to review approaches from their own field and identify best practice that could be emulated. The existing guidance was felt to be too generic and vague by some, so enhancing this with discipline-specific examples and pointers would help to contextualise the questions and give approaches to follow. Indeed, a few people commented that a more interactive template that provided recommendations and dropdown options instead of so much free text would be preferred.

Currently, [a new RDA working group](#) is developing **common standards for DMPs**. The aim is to define a common information model and specify access mechanisms that make DMPs machine-actionable. The outputs of this working group will help in making systems interoperable and will allow for automatic exchange, integration, and validation of information provided in DMPs. Although these aims were a lower priority for survey users, ensuring the data gathered in DMPs is put to good use will help to derive more value for researchers and prevent it being perceived as an administrative exercise. The PIDs required by the EC could be used to link up information held in DMPs with other systems, improving data discoverability and assisting in monitoring and reporting.



Recommendations: Restructuring the DMP template

The chief recommendation we make is to consider **restructuring the DMP template** to make it easier to complete. The concept of FAIR works well, but ordering questions in this way causes overlap and redundancy. Questions could be more usefully grouped thematically or by key activities (e.g. collecting or generating data, developing and documenting metadata, and data sharing). This would provide a more logical arrangement and avoid repetition, while still addressing the desire to result in FAIR data.

Information to evidence and assess the FAIRness of data could be collected in the final version of the DMP (e.g. PID, repository details, metadata), ideally in a form that allows for **automated evaluation**. This could be done by determining key aspects to evaluate and reflect on data discoverability and its monitoring and reporting with specific fields in the DMP template. Providing a machine-actionable DMP template will help to realise automated evaluation of DMPs.

Providing **dropdown options** would further enhance the template's user-friendliness. This may be difficult to achieve, but a move towards more structured responses or amendment of the template in order that irrelevant sections / questions can be skipped should be investigated. It would also help to differentiate **which questions should be addressed when**, with a minimal set of questions that are required at month 6 and more detailed prompts during the project as the approach is developed. This is not to preclude projects providing richer information earlier, but could help to avoid researchers unfamiliar with the process from feeling overwhelmed by seeing so many questions and unsure which are of most importance and relevance.

Recommendations: Tailored guidance, examples & terminology

The survey brought to the fore that respondents want **much more tailored guidance** and **discipline-specific examples** to help them apply the DMP questions to their context. We therefore recommend that the EC works in collaboration with the ESFRIs and disciplinary groups to provide domain-level guidance. The existing Science Europe domain protocols initiative could be built upon. Extra guidance should also be provided for larger, more complex projects with multiple work packages, data outputs and stakeholders, as they often found the DMP complicated to complete.

Respondents demonstrated that there is a willingness to **openly publish DMPs** (48% yes [Q12]). Encouragement of this practice by the EC would be beneficial, as early release would allow communities to learn about research going on elsewhere and promote collaboration and sharing of outputs. The example DMPs that are shared should be collated in a **DMP registry** to improve discoverability and reuse of the content. We also recommend that the EC shares a small collection of 'approved' DMPs to give an indication of good practice in different research areas as a guide. Collaboration with EC-funded e-infrastructure projects, ESFRIs and external reviewers would be beneficial here.

We advise to **simplify the terminology** used in the template in a way that both researchers and support staff can relate to it. Where possible, overtly technical terms should be replaced with more explanatory wording. Where terms that have caused confusion are retained, such as interoperability, metadata and ontologies, attention should be drawn to them by providing a glossary and/or examples in the template itself. This will help DMP writers to gain a better understanding of the terminology used in data management.



Recommendations: Support on costing and rewards

Clear examples are needed on the types of RDM costs that are eligible to be included in proposals. We recommend that the EC provides **worked examples and guidelines on costing** (e.g. [Wellcome guidance](#) in section 4 and [LCRDM cost guide](#)). This would help projects to understand the kind of activities involved, what time and effort may be required for these and any direct costs or charges that could be applied.

There should also be a requirement for all projects to consider RDM activities at the application stage to prompt them to include associated costs. Researchers, and among them especially early-career researchers, should be adequately rewarded for spending time on implementing their DMP and demonstrating good data management and sharing practices. This should form part of the EC's reporting and review mechanisms.

Recommendations: Clarifying the review process

There were also a number of comments in the survey seeking clarity on **how the DMPs will be reviewed and by whom**. Respondents did not always understand the rationale for each question and how it would be used or assessed by the reviewer. This caused some to stick to the template or perceive questions as mandatory despite feeling certain aspects were not relevant. There was also confusion about the RDM section in the grant application and anecdotal evidence about projects that opt-out or do not provide a DMP being evaluated unfavourably.

Answers to question 9 confirm that more detailed information about the review process is needed. Research support staff would like to know whether any common standards or review criteria are being followed. We recommend endorsing the internal assessment framework developed by the Research Executive Agency (REA) and providing **review guidelines** so all project officers and external experts brought in to review DMPs follow the same approach. **Publishing the assessment framework** and a statement on the processes being followed would also be beneficial to the wider community for transparency.

Although the Commission states that the DMP template is only a guide, not mandatory, there is a tension with this being used as a review framework since DMPs that don't conform to that structure may be poorly rated. Moreover some users have been informed they are using the wrong template, or are confused about whether they should answer the set of questions in the 'Template' or the 'Summary Table' of the [official guidelines](#). If the template is to remain optional, it would be helpful to highlight a subset of concepts or aspects that must always be addressed so researchers know what has to be covered in their DMP, irrespective of what format they provide this in.

The Open Data Pilot has already brought significant benefits and driven the development of RDM support and infrastructure in many organisations across Europe. We applaud the changes implemented so far in response to community feedback - the widening of participation so all work programmes are included by default, and the transition to FAIR to emphasise the importance of good data management practices, not just open data. Although several further pointers for improvement have been made in survey responses, it is clear that many aspects of the approach are working and being well-received. Some adjustments to the template and provision of disciplinary guidance and examples will simply enrich the DMPs being provided and provide a firmer footing for the transition from pilot to policy.



Outcomes per question

In this section, short texts supplemented by graphs or word clouds summarise the outcomes per question. These are followed by a selection of respondents' quotes. An overview of and insight into the outcomes are summarised in an infographic in Appendix I.

Question 1 - Role of participants

For the purpose of the survey we distinguish two roles. If you belong to both groups, please check the appropriate box and use the open text fields for extra information when your experience in these roles varies.

[289 respondents, 132 comments. Answers: "DMP writers" 50%, "DMP support staff" 61%]

Division of roles of respondents

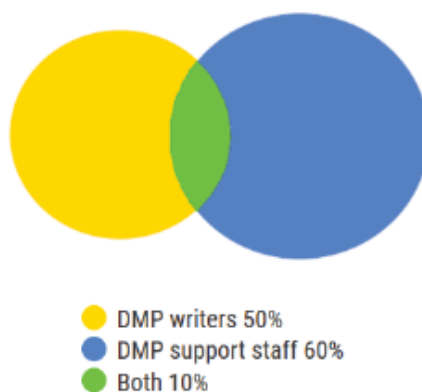


Figure 002 - Q1: role of respondents

From the 289 respondents, 145 stated that they are "DMP writers", while 176 indicated that they fulfil the role of "DMP support staff". 132 respondents, furthermore, commented on their roles as they could not classify themselves among the two categorisations given. The majority of the respondents who commented on their roles (79/132) are librarians and/or project managers or project coordinators. Three of them give training; three work for organisations or companies offering expertise and software to establish DMPs. 12 respondents state that they take up the role of both researcher and support staff. 2 respondents, finally, are researchers in the field of Data Management.



Respondents commented as follows:

"I review draft plans, providing guidance on answering the questions, suggesting 'boilerplate' text that describes institutional facilities and provide example text where helpful."

"I run a business that consults about, creates, and executes FAIR data management plans."

"Managing the full research data lifecycle. PhD student in Data Management."

"I am the deputy project coordinator, dissemination work package leader and researcher within a Horizon 2020 project. I wrote our project DMP, with support from my university's central data management team."

Question 2 - Name and affiliation

Please note your name and affiliation. We do not publish combinations of names and individual answers.

This question is removed due to privacy regulations.

Question 3 - Reaction to H2020 DMP process

Did you find the process of writing / supporting a H2020 DMP positive or negative?

[189 respondents, 119 comments. Answers: "Positive" 60%, "Negative" 16%, "Not applicable" 24%]

Overall experience with writing/supporting a H2020 DMP

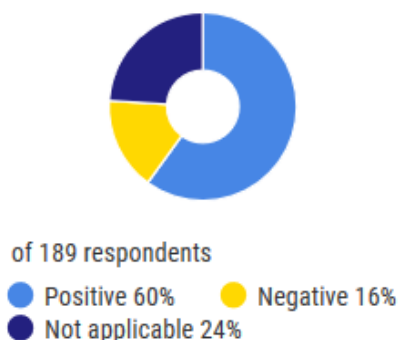


Figure 003 - Q3: reaction to H2020 DMP process

60% (114/189) of the respondents indicated that they found the process of writing / supporting a H2020 DMP positive, while 16% (30/189) experienced it as negative. For 24% (45/189) this question was not applicable. Respondents' comments indicate that, overall, they felt that the process of writing a DMP made them reflect on how to deal with data. It also helped to clarify questions about data management and think about the structuring of data at an early stage of the project. A minority of respondents were negative about the highly administrative efforts to set up a DMP, not least because the EC's template uses too much jargon unfamiliar to researchers. For a detailed list of unclear terms see Question 5.

Respondents commented as follows:

"Both positive and negative: it helped us reflect on potential issues and decide how to address these as a project, and also it was frustrating and felt like an administrative exercise that was very long and cumbersome; too specific in some areas and too vague / unclear in others."

"I cannot say it was negative or positive. It stimulated interesting discussions among the leading scientists, who had very different view what a data management plan was about in the first place. Some intended to write legacy, but could be convinced that this was not about developing new data policy, because the project only re-uses data from many different sources. This specialty was also the reason that the due date of a first data management plan after 6 months was felt far too early for this particular project. At this stage, it was not sufficiently settled what data sources would be used and how the data should be handled, stored, and made accessible to external users. Therefore, to some extent, there was unnecessary effort with the risk of writing or agreeing things that would become obsolete very quickly. Because of "theoretical" base in the beginning, some scientists were tempted to write more a proposal than data management plan."

"It helped to structure the data of the project, not only for project partners but also for other people who can reuse some data."

Question 4 - Experience with H2020 DMP template

To what extent do the following statements represent your experience of using the H2020 template?

[173 respondents. For percentages see table below]



	AGREE	NEITHER AGREE NOR DISAGREE	DISAGREE	TOTAL
It is a very useful template.	61% 105	23% 40	15% 26	171
The structure is clear and it groups questions into helpful categories.	65% 111	19% 33	15% 26	170
I like the level of guidance and contextual information that is provided.	45% 77	30% 51	25% 42	170
There are too many free text answers. I need more drop-down options.	44% 74	31% 52	25% 43	169
There are too many questions!	31% 53	38% 64	31% 52	169
Many questions are irrelevant. I'm not sure why they are being asked.	25% 42	36% 61	39% 65	168
I don't understand what FAIR means.	10% 17	16% 28	74% 125	170

Figure 004 - Q4: experience of H2020 DMP template

The template and grouping of questions was found useful, the questions relevant and clear. Despite the fact that for a small majority of respondents the level of guidance and contextual information is appropriate, respondents indicate that they would appreciate more drop-down questions. The concept of FAIR is clear to the majority of respondents, although responses to questions 5-8 make clear that there are questions about implementing the FAIR data principles.

Question 5 - Unknown, unclear or confusing terminology

Terminology - please note any terms that are unknown, unclear or confusing:

[80 respondents, 80 comments, which often contain several terms that are unclear or unknown]

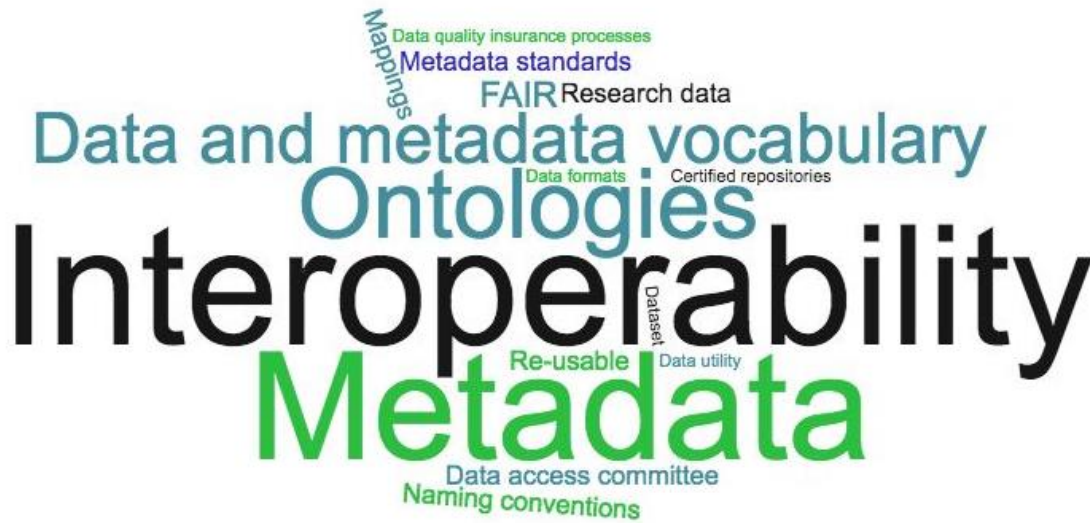


Figure 005 - Q5: unknown, unclear or confusing terminology

From the 80 respondents, only 13 indicated that the terminology was clear. Among those who listed unclear terms “interoperability”, “ontologies”, and “metadata” were mentioned most frequently. In total, respondents mentioned as many as 40 terms which were unclear to them, among which terminology concerning “data” in the broader sense occurred 6 times. Some respondents also commented that they would appreciate brief explanatory texts for every question in the template (see also question 6).

Respondents commented as follows:

“Maybe it should be made clearer what is meant by metadata. Usually, I tend to consider as metadata all the material attached to data, like codebooks and technical annex, as well as all the information on sampling, geographical coverage and so on. However, it would be better to include in the template an explaining-box with examples of metadata.”

“With the help from Swedish National Data Service we could clarify many questions. Without this help we would not be able to finish the DMP.”

Researchers are not familiar with the following terms/phrases : Metadata, standards for metadata/data, ontologies, mapping with ontologies, interoperability, All the ICT jargon

“Although there is an attempt to define “interoperability” I found the explanation incomplete, too vague.”



Question 6 - Confusing, inconsistent or redundant questions

Questions - please note any questions that are confusing, inconsistent, or redundant:

[82 respondents, 82 comments]

23 questions were mentioned as being unclear and/or problematic by more than one respondent. Among the unclear questions the following six were mentioned most frequently:

- “What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.”
- “If there are restrictions on use, how will access be provided?”
- “Is there a need for a data access committee?”
- “Are data quality assurance processes described?”
- “Do you provide clear version numbers?”
- “Are there well described conditions for access (i.e. a machine readable license)?”

Some of the respondents would prefer to have questions grouped more efficiently and would like to have structured vocabularies tailored to their fields at their disposal. The question “What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?” was mentioned as being redundant and/or irrelevant most often, namely by 4 respondents.

Respondents commented as follows:

“To researchers it looks like there are repetitions of questions about the Metadata as they appear under 2.1, 2.2 and 2.3. This is due to the fact that the subsections are based on the FAIR principles rather than overall defined categories such as: data description, data collection, data management (storage, security), metadata, data sharing.”

“There appears to be some duplication, reflecting questions that ask similar but not identical questions. Some effort is needed to resolve this as users don't want to re-write similar information multiple times.”

“Are data quality assurance processes described? The DMP itself is a quality model and making the data interoperable and accessible is part of a quality process. So, this question may be a bit ambiguous and require examples to be properly understood.”

"Rather than confusing, i don't like how questions are arranged along the text since they don't follow a logical approach. Having to explain how to make data findable before talking of the repositories chosen to make data open is difficult and forces you to anticipate things to be detailed later. i would like to have the guidance questions not being forced into a specific structure"

Question 7 - Missing or potential guidance

Guidance - please note any guidance that is missing or could be improved

[81 respondents, 81 comments]



Figure 006 - Q7: missing or potential guidance

Only 17 of the 81 respondents indicated that the guidance was adequate. The 5 most frequently mentioned desired additions or clarifications with regards to the guidelines were:

- best practices, examples or a reference DMP
- subject-specific templates
- where to find more information or specialised guidelines for a certain field
- guidance is too complicated, technical, vague, or generic
- information about costs

The most requested addition, that of inclusion of examples, returns in many different forms. Example DMPs, example answers for specific questions, discipline-specific examples as well as examples on how DMPs are reviewed are all mentioned. The three first topics - examples, subject-specific information and the need for more specialised guidance - can be seen as a demand for information on how to complete a DMP in general. It is clear though that respondents would like to see the guidelines more tailored to the specific research context. Interoperability was most requested as a specific topic where examples would be appreciated.

Also, respondents find it difficult to estimate costs for data management and storage. Further concerns deal with the level of expertise expected from researchers to understand and interpret the guidance and ambiguities about what level of detail is expected in each question. Also, some remarks about the structure of the template were made in this context.

Respondents commented as follows:

"There have to be examples of correct DMPs with explanations why they are good. There should be also "reviewers guidelines." Currently it is not clear who reviews DMPs and what is expected from them."

"I didn't find any guidance! I found examples of other DMPs on the www."

"Guidances are too general. Disciplinary guidances and examples would be useful."

"Guidance should have example answers or links to external resources."

"A more clear, consistent and transparent template. Not just a list of open questions. Also a clear (tabular) structure would help."

"This template allows for too much freedom. So unless you have access to a FAIR expert most researchers I know do not really know where to begin. I would like to enable researchers to use a structured tool to fill in this template, and promote an interoperable format to exchange such templates (i.e. to enable professionals to fill it in together with researchers, enabling reuse of best practices on coarse grained level). E.g. at the Dutch research foundation (ZonMw) we use existing tools from the data curation center UK."



Question 8 - Topics that are missing or should be removed

Coverage - please note any topics that are missing or ones that should be removed:

[64 respondents, 64 comments]

Of the 64 respondents to this question one third had no further recommendations. The remaining answers cover a broad area of issues which can be categorized in ten missing topics:

- **Discipline-specific approach.**
- **Access:** information about closed data and levels of restriction.
- **Data quality assurance:** too little attention to data testing, quality and validation procedures, quality assurance, measuring uncertainty for data and risk management
- **Metadata and documentation:** questions how to adequately document uncommon data and data types were posed, e.g. data for which new standards are created, interdisciplinary data. and documentation other than structured metadata.
- **Software:** a lack of questions addressing preservation of software, code, and coding standards.
- **Reproducibility:** cases cited involved data processing, lineage, provenance, and linked data, the capture of models, algorithms, and associated metadata, the long-term aspect of data provision and potential reprocessing.
- **Costs:** Information could be added about resource limitation and resource management on collaborative environments.
- **Storage:** the importance of a good repository is not stressed enough in terms of long-term storage and possibilities in terms of metadata and licensing. Objectives and strategies to maximise exploitation of results are not fully addressed.
- **Legal issues:** there were some remarks on adding questions or guidelines on ownership of data, privacy policy, ensuring ownership, or right to publish, international legal issues concerning archive data, conflict of interest, and data anonymisation.
- **The format of the template:** There were several suggestions to supplement or amend the current format of the DMP template, namely a clear definition of terms, a timeline, following up on or integrating practical implementation of the DMP, group questions by topic, drop down menus and more closed questions.

There were only a few suggestions of topics that could be removed and they often conflict with previous requests to add information on certain topics. Concerning the topics requested to be added or extended we are able to state that most of them are already covered by the current DMP template to some extent. Often, it is more a question to what level of detail the topic should be covered and specific questions eligible for peculiar cases.



Respondents commented as follows:

"I think this would probably vary according to individual projects. It is possible that a 'one-size-fits-all' approach is not practical."

"Consideration of problems caused by the use of international research data in the humanities/social sciences : e.g. legal issues concerning archive data from China used in research projects in Germany."

"The template is generally seen as very elaborate and time intensive. It would be good to group questions, reduce the number of questions or replace 'open' questions with closed ones. I also would be in favour of a more 'practical' instead of 'theoretical' DMP. e.g. provide a link to the dataset in a repository rather than asking what repository, what dataset, under which licence....it would do more to provide a cultural shift towards data management than writing a DMP, seen by many as an administrative burden even when in favour of Open Science."

Question 9 - Issues encountered following EC guidelines

Process - please note any issues encountered following the EC guidelines e.g. knowing when a DMP is due, how it will be reviewed, how to include costs etc:

[68 respondents, 68 comments]

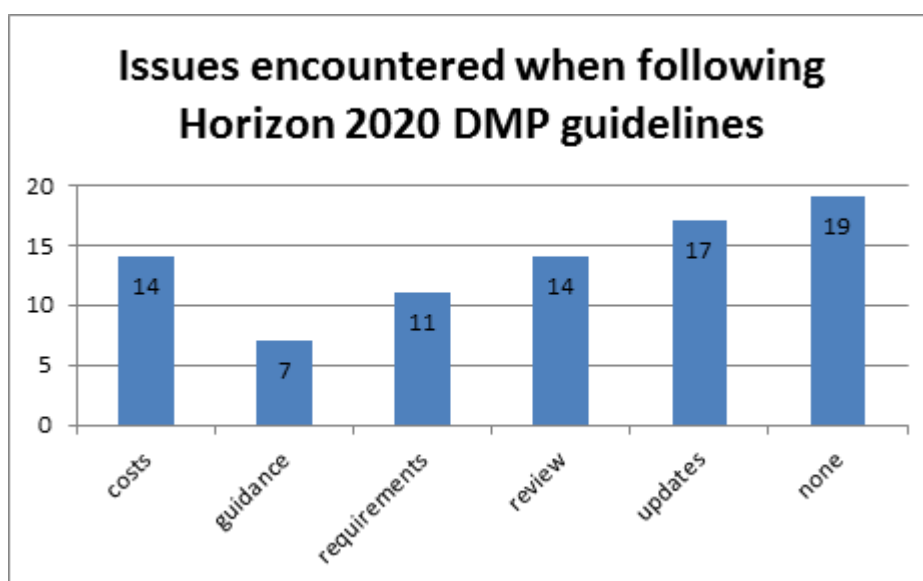


Figure 007 - Q9: issues encountered following EC guidelines

There were 68 responses to this question, raising 82 discrete comments on different themes. Only 23% of the comments (19/82) reported no issues in following the EC guidelines and felt the process was clear. The majority of comments (77% - 63/82), however, reported issues. These were predominantly related to the process of updating the DMP during the project, estimating and including associated costs, and understanding how DMPs would be assessed and monitored.

17 comments related to the process of updating the DMP. Questions were raised about what information should be provided when; indicating that a more structured template that increases the level of detail as the project matures would be useful. Technical issues were also reported about the portal as existing DMP deliverables need to be rejected for new versions to be submitted, and the versioning process isn't clear. Costs were a primary concern in terms of implementation. Participants did not know how to include costs, were unsure which items were eligible, and struggled to make accurate estimates. The other key topic where respondents sought more information was in terms of the Commission's DMP review practices. It was not clear how DMPs are being reviewed and by whom, or whether any common standards or review criteria were being followed.

Respondents commented as follows:

"Which parts should be set up already in the beginning and what can be added later? Would be helpful to indicate "needed" and "nice to have" for the starting point or for different stages."

"Processes around the assessment of the content are unclear to me."

"As stated before, clarification of what is an eligible cost regarding RDM is urgently needed, for researchers and also for the RDM support teams at the universities. For example:

- *If a project needs to buy extra storage from the own institution, is that cost eligible? Or only external storage systems are eligible?*
- *Database curators/admins could be financed by EC projects?*
- *Are licenses for Electronic lab Notebooks (ELNs) for experimental data documentation eligible costs?*
- *Are there differences in RDM eligible costs between consortium projects and IF or ERC, for example? If there are differences, it would be best if the EC provided examples for all the type of calls.*
- *When to include those costs? At the proposal stage? Is it also possible while the project is running?"*

Question 10 - Suggestions on the DMP template

Other - any other issues or suggestions regarding the template you would like to make:

[48 respondents, 48 comments]

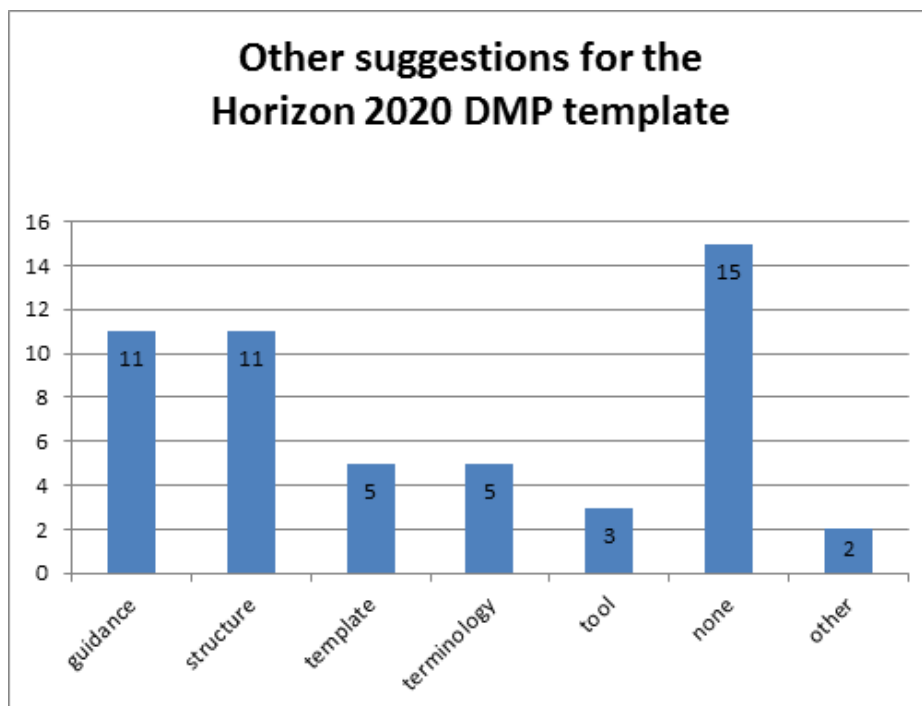


Figure 008 - Q10: suggestions on the DMP template

There were 48 responses to this question, which raised 52 discrete comments on different themes. 15 respondents had nothing to add here (15/52). Of the suggestions provided, most were related to the need for additional guidance, a more structured template or terminology issues. The eleven comments related to guidance were primarily seeking examples of good DMPs or discipline-specific guidance and options. Eleven comments were also made about the template structure. Most of these focused on making the template shorter, providing checkboxes, optional questions or clear pathways through so it was more tailored to each project's context. It was suggested to use two or three different templates for initial, mid-term and final DMP to better reflect the evolving process of DMP writing. We don't support the latter suggestion. On the contrary, this very practice during the initial stages of the H2020 Open Research Data Pilot was perceived as impractical and going against the grain of a living documents; early 2016 the OpenAIRE and EUDAT projects have recommended that there should be just a single DMP template, with clear markings as to what should be answered initially and what should be addressed later.

FAIR data, while felt useful as an overarching concept, was not considered to be a helpful way to structure the template. This is borne out in responses to earlier questions that note significant overlap in the FAIR criteria and confusion about what to write where.



Respondents commented as follows:

“Perhaps it would be interesting to include a base list of data typologies common to all projects, with name, data type, extension, associated security type, and traceability.”

“What's the purpose of the SUMMARY TABLE. Just a summary of the template or a table to be filled in by the DMP writer?”

“Less free text fields, and more boxes to tick (especially for the initial DMP).”

“We do not think the one-dimensional nature of the conceptual model underlying the DMP is workable. There are multiple linkages between funding organisations, researchers, outputs, grants, etc. that are not captured by the template as it stands. As an absolute minimum, each enumerated output should have a PID so that it can be explicitly linked to a deposit action later in the life of the research project.”

Question 11 - Priorities for a DMP template or tool

In an ideal world, what would you like a DMP template or associated tool to do for you? Please rank your top 5 priorities below (one response per column):

[138 respondents, 15 comments]



Priorities ranged by weighted average:

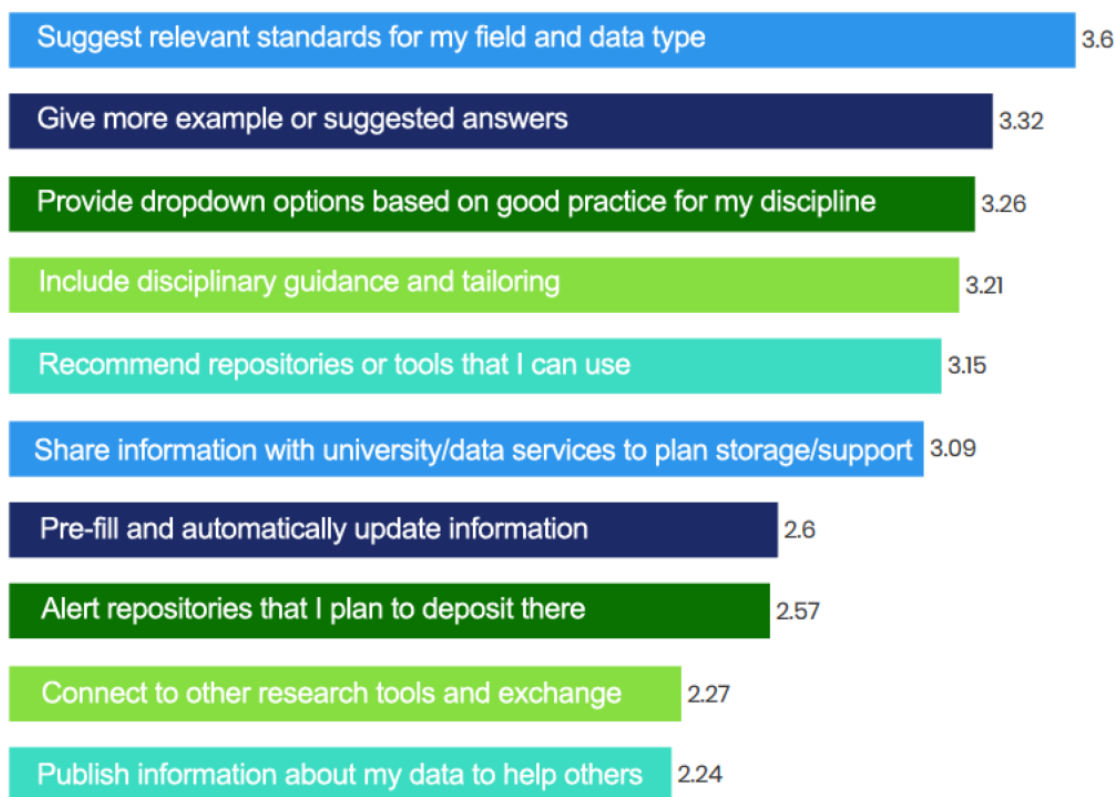


Figure 009 - Q11: priorities for a DMP template or tool

This question presented respondents with 10 suggested improvements and asked them to rank these to indicate their top five priorities. The graphic above shows the results in order of preference, based on the weighted averages. The most desired feature is to suggest relevant standards for the user's field and data type. This mirrors the findings of an earlier USA study that analysed 500 DMPs to better target institutional support ([Wells Parham et al., 2016](#)). This found that 81% of the DMPs failed to specify a metadata standard, as there was a low awareness of relevant options across many disciplines. As indicated in earlier questions, respondents want much more tailored guidance and discipline-specific examples to help them apply the DMP questions to their context. Recommendations of tools and repositories that can be used also came high on the list.

Respondents could see a benefit to sharing information with university or data services to plan storage allocations and support, but some of the other machine-actionable DMP use cases, such as pre-filling information and exchanging data across tools were lower priority, indicating that the most pressing need for support is additional guidelines, recommendations and worked examples.

Respondents commented as follows:

"Gather good practices from other DMP and add this practices in the template."



"It would have been helpful to see another project's approved DMP. All researchers that I have been in contact with had only DMPs that were not approved."

"It would be great if it all worked in a linked data way to reduce the amount of double keying or copy and pasting of data."

"Unless you can be very confident about that the system recommendations are in line with requirements of the depositing archive I would be extremely careful about doing this."

Question 12 - Openly publishing DMPs

Would you openly publish your DMP?

[134 respondents, 43 comments. Answers: "Yes" 48%, "No" 20%, "Please comment" 32%]

Percentages of respondents on whether to openly publish DMPs

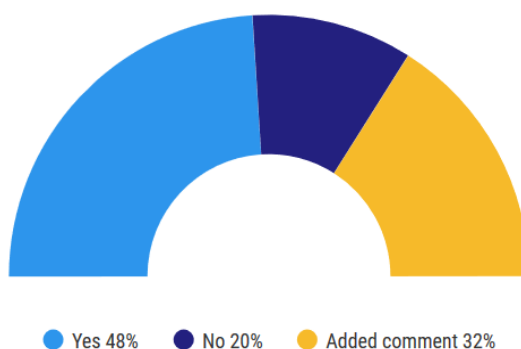


Figure 010 - Q12: openly publishing DMPs

48% of the respondents (64/134) would openly publish their DMP, whereas 20% (27/134) would not. More than half of the free text answers can be summarised as "Yes, I would, depending on...", for instance depending on the confidentiality of the information in the DMP (if any) or "probably once the project is finished". Respondents also wish to reach a level of confidence before publishing their DMP openly: "as soon as we have a good workflow for RDM." Others indicate that the DMP cannot be shared, for instance when project partners work in circumstances that might be dangerous. With minor exceptions, RDM supporters tend to leave the decision to the researchers.

Given the strong wish for example DMPs we recommend to encourage projects to openly publish their DMP and to plan accordingly in the project proposal deliverable list.



Respondents commented as follows:

“As a librarian I recommend DMP publishing to our scientists and would welcome the EC to recommend or demand this for EC funded projects.”

“We are treating sensitive information in the DMP as regards how our partners will operate in difficult contexts, for instance in Egypt.”

Question 13 - Resources to include in disciplinary guidance

The European Commission plans to add disciplinary guidance to its generic H2020 DMP template. What standards, tools, databases, repositories or other support and best practices should be referenced for your discipline?

[69 respondents, 69 comments]

69 respondents offered a considerable amount of suggestions for the use of standards, tools, databases, repositories, and other support throughout various disciplines. Zenodo was mentioned most frequently (9 times), followed by Github (4 times), Copernicus Data Hub (3 times), and Biosharing (3 times). 8 respondents pointed to the fact that their disciplines first need to clarify and define what exactly data is, or that standards are not established yet. Comments were also made on the guidance for repositories. It was felt that a list of approved repositories or repositories provided by the EC are needed. One respondent commented that currently very little guidance on licensing is available. These responses indicate that disciplinary guidance would indeed be very valuable.

Respondents commented as follows:

“In the humanities, the discussion on DMPs and first on what research data are in the respective discipline is just at the beginning. Therefore, no answer can be given to the question right now. There needs to be more consideration of the affordances and needs of the humanities (esp. area studies subjects with research data in non-latin scripts) and more time for discussions about it. Humanities experts are not that well represented in RDM expert groups.”

“Unfortunately, this is exactly where I perceive the problem to be - I am not aware of standards etc for my discipline and would ideally like the EC to provide this information to me, rather than the other way around.”



“Need to be very careful <when> correct domain guidance is given. The bespoke requirement of individual of projects need to be discussed with domain data experts.”

“biosharing.org is kept up to date. It is better if resources like that are referenced than trying to include this information in the template.”

Question 14 - Suggestions to EC on DMP approach

What are the main suggestions you would give to the European Commission about its approach to DMPs in the Open Data pilot?

[75 respondents, 75 comments. 26%]



Figure 011 - Q14: suggestion on DMP approach



There were 75 free text comments as to suggestions about the approach of the EC regarding the DMP template, most listed more than one suggestion. The six most frequently mentioned suggestions were:

- A more customised and/or discipline-specific template and guidance
- Include (interoperable) tools and link to existing services, make DMPs machine readable
- Provide examples or share approved DMPs
- Simplify the template and reduce the effort for researchers completing the DMP template
- Provide incentives and make the added value of data management clear
- Involve researchers, universities, data centres and member states when updating guidance or the DMP template

A recurring suggestion is the need for a more discipline-specific approach whether in the template itself or in the accompanied guidelines. There were several suggestions asking for more openness, not only in providing examples of successful H2020 DMPs but also in regard to the reviewing process of the EC. Providing feedback on submitted DMPs or compiling a checklist of the vital criteria used to review H2020 DMPs would aid research support staff in guiding projects to successfully comply with the requirements.

Further education in the form of workshops around the requirements of the H2020 DMP template is welcomed too. Information about possible costs and how to fund them is reported missing. Finally, suggestions were made to take a critical look at the timeline of the DMP requirements and when in the research lifecycle these requirements are best placed.

Respondents commented as follows:

“Talk to researchers and ask them whether the DMP requirements are realistic. Involve researchers when developing disciplinary guidance.”

“Focus on technical exchange format for these plans so they can be shared between tools (i.e. make the plans themselves FAIR)”

“It is good to have a single template for all disciplines, but disciplinary (or data-type-specific) examples, guidance and exemplars should be available and linked. Re-emphasize the need to integrate practical data management and curation into real-world, day-to-day research workflows rather than an afterthought.”

“Think about incentives not just obligations!”

“Connected with external tools (default repositories, etc.)”



Question 15 and Question 16 - Providing feedback on DMPs

Q15: *Have you given feedback to (other) DMP writers on their draft DMP?*

[139 respondents, 139 comments. Answers: "Yes" 37,41%, "No" 62,59%]

Percentages of respondents giving feedback on draft DMPs:

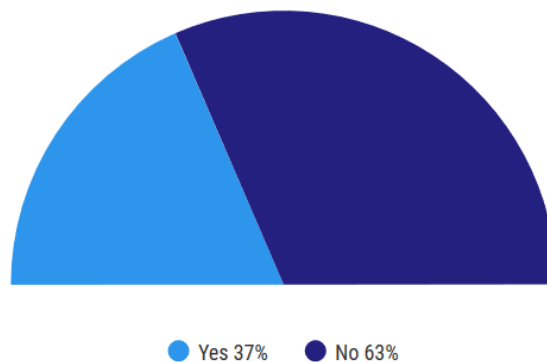


Figure 012 - Q15: respondents providing feedback on DMPs

Q16: *What kind of feedback did you give (multiple choice)?*

[49 respondents, 13 comments. For answer percentages see graph below]

Graph for Question 16

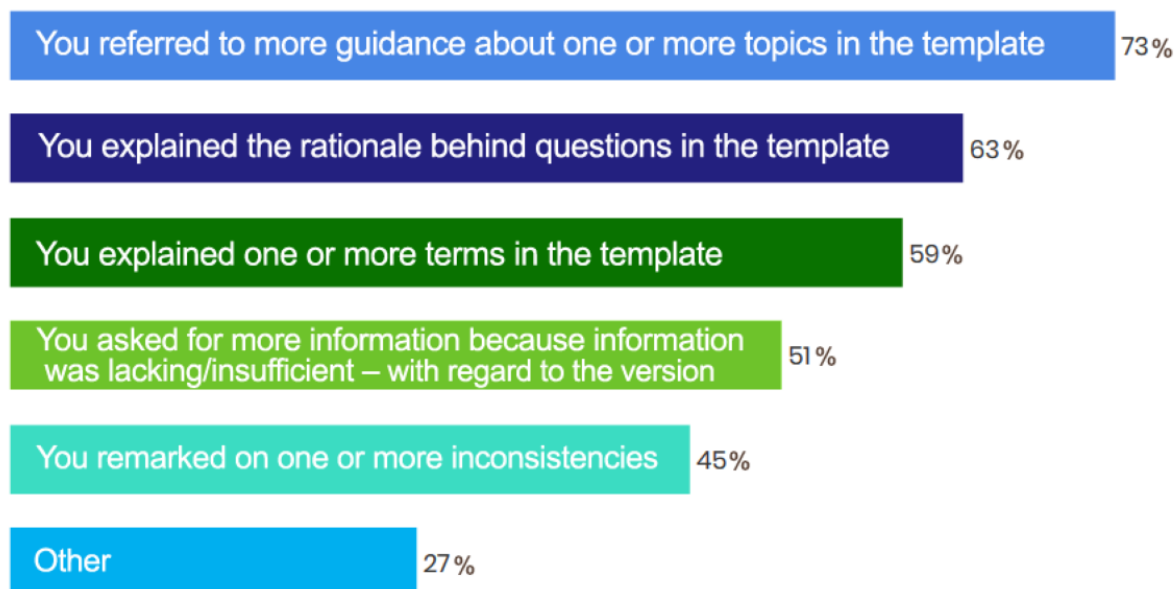


Figure 013 - Q16: kinds of feedback on (draft) DMPs

37% of the respondents (52/139) have given feedback on draft DMPs. Most frequently (36/49), they referred to more guidance about one or more topics in the template and explained the rationale behind one or more questions (31/49). (The same question in the 2017 [LCRDM survey about DMP experiences](#) shows even higher percentages of feedback. In the same order as in the figure above, the LCRDM percentages are 81%, 85%, 85%, 81%, 62%, and 36% other feedback. Respondents of the LCRDM survey reviewed DMPs for various funders, and it is not possible to relate any findings to a particular funder or DMP template.)

Free text answers to question 16 indicate that they explained concepts, referred to tools, systems, and examples, offered sample text, and assessed the DMP. They also encouraged researchers to integrate data management planning in all relevant work packages and stressed compliance with institutional and legal norms. One respondent created a bespoke DMP appropriate to the project, which also incorporated all elements from the template.

Respondents commented as follows:

“In some cases I had to look together with researchers into their publications or WPs description of a project in order to help them identify the type of data they are collecting in order to include them in the DMP. This exercise also helped them to identify which data might have a re-use potential and which ones were submitted to special policies/regulations.”



“Advised on data storage and sharing systems available at University.”

Question 17 - How H2020 DMP support process differs

How did the process of reviewing and supporting researchers with a H2020 DMP differ from assisting them to write plans for other funders or contexts?

[45 respondents, 10 comments. For percentages see graph below]

Percentages of outcomes on how the process of reviewing H2020 DMPs differed from other funders' DMPs.

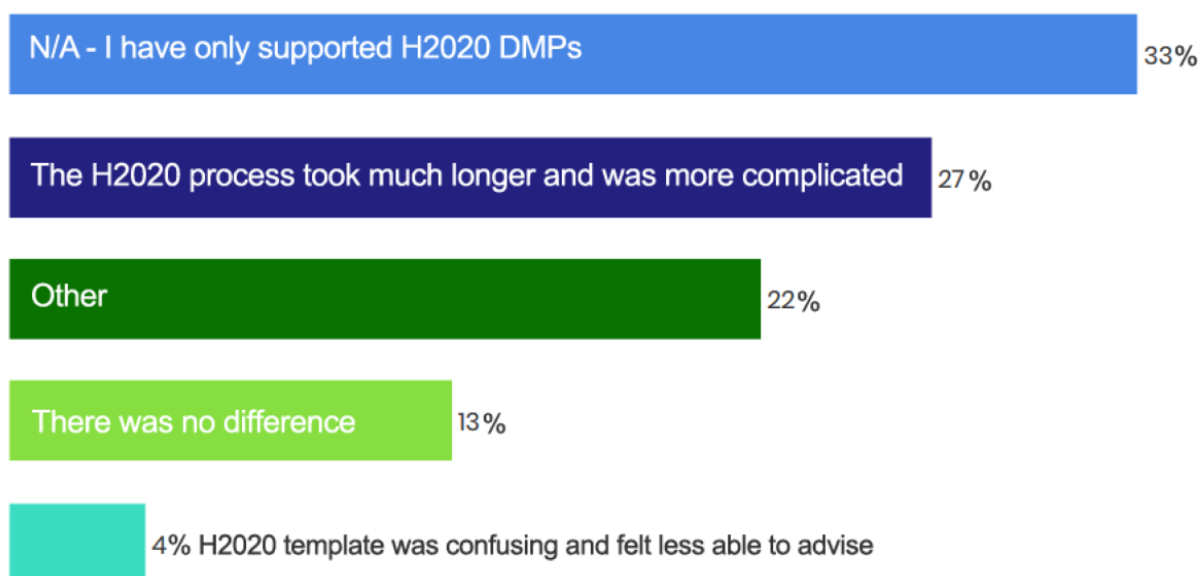


Figure 014 - Q17: comparison H2020 DMP and other DMP support processes

33% of the respondents (15/45) indicated that they have only supported H2020 DMPs. 13% (6/45) experienced no difference between supporting researchers with a H2020 DMP and other funders or contexts, however, 27% (12/45) state that the process took much longer and was more complicated than assisting researchers to write plans for other funders or contexts. This may have to do with the fact that DM planning and support is still new - only 6 respondents provided text on the comparison, so this number of free-text responses is too low to draw conclusions. No-one chose the option that the H2020 process is quicker and easier, although one respondent remarked that it was quicker but provided less structure.



Respondents commented as follows:

“The process is not yet standard: the quality of the plans I have seen so far is so different that it is impossible to give averages.”

“We hardly ever support H2020 DMPs. Initially, it is not clear that the DMPs are not mandatory.”

“The H2020 process took a little longer (requested more information)”

“Similar level of work required, but made more enjoyable and effective because researchers are dedicated and working towards submission, rather than just making one because I tell them EPSRC expect them to have one in place.”

Question 18 - Time needed to review H2020 DMP

How much time do you need to review a draft H2020 DMP on average?

[50 respondents. For percentages see diagram below]

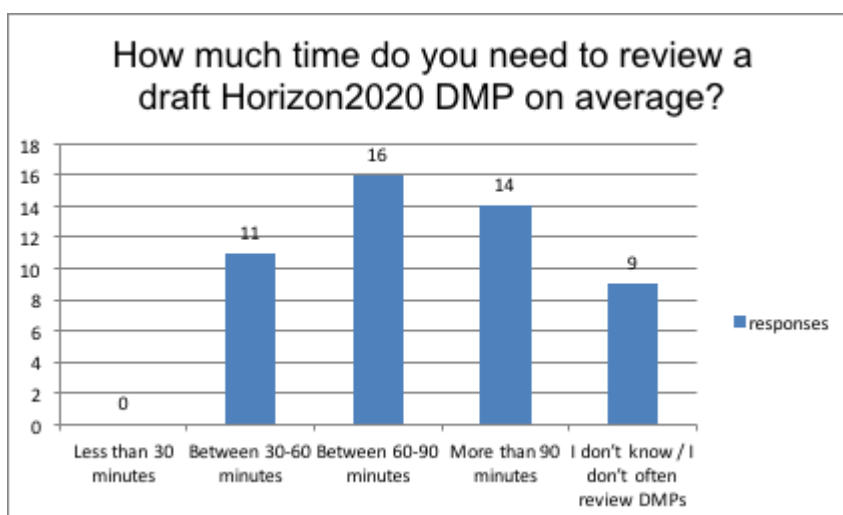


Figure 015 - Q18: time needed to review H2020 DMPs

22% of the respondents (11/50) indicate that they need between 30-60 minutes to review a draft H2020 DMP. 32% need between 60 and 90 minutes, another 28% need more than 90 minutes. The remaining 9 respondents selected the option “I don’t know/ I don’t often review DMPs”.

The same question in the [LCRDM survey about DMP experiences](#) shows a similar curve but more to the left, i.e. reporting a shorter duration. For the LCRDM respondents, who reviewed DMPs for various funders, based on various templates, the numbers can be seen in the next figure. The LCRDM survey doesn’t allow for linking the indicated effort to a particular template, but these different outcomes suggest that reviewing a H2020 draft DMP takes more time than reviewing DMPs based on other templates.

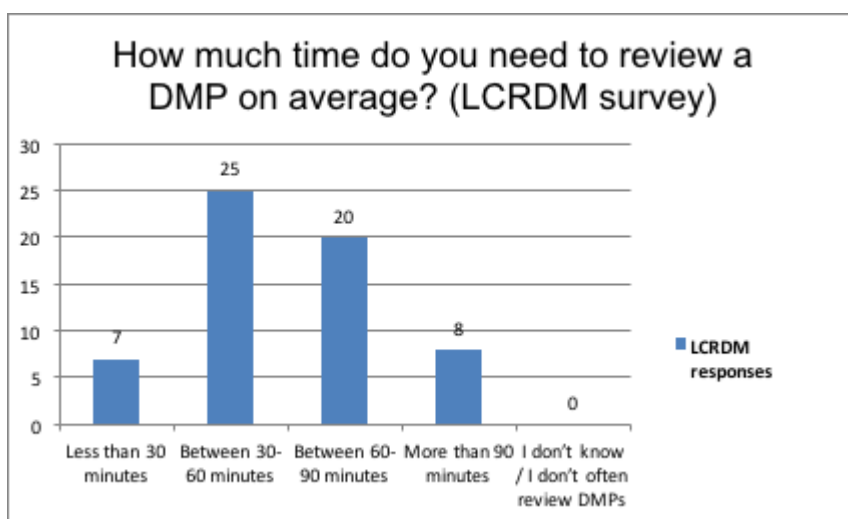


Figure 016 - LCRDM survey: time needed to review DMPs

Question 19 - Other recommendations on FAIR data management

Do you have any other recommendations about the Guidelines for FAIR Data Management?

[39 respondents, 39 comments]

There were 39 comments to this question, half of which indicating that there was nothing further to say (20/39). The recommendations made by the other 19 respondents concern primarily the need for more or better guidelines, clarifications, a checklist, and good practices. For instance “It would be great if the EU also provided an easy-to-read checklist for DMPs to assist data managers with reviewing submitted DMPs.” Another respondent expressed the need to clarify and consider what a DMP is for:

“Is it for a reviewer to appraise the project, the project officer to monitor the project, to inform an archive about data or to help the project do data management planning? The DMP should not attempt to fulfill all these functions.”

“Some of the most useful guidance is actually in the “Open access to publications and research data in Horizon 2020” document. In particular, the categorisation of “underlying data” and “other data”, and the steps outlined under “What are the requirements of the Open Research Data Pilot?”. Although this document is linked to early in the “Guidelines for FAIR Data Management” document, I think it is very likely that researchers miss one or other of the documents (they already have to deal with so much documentation!). The documents should be more explicitly linked (i.e. more specific references from one to the other at various points) or preferably merged to provide a single overview and to remove any inconsistencies or differences in emphasis.”

7 recommendations - again - relate to the FAIR data principles, and they are rather critical:

“There is too much overlap between F, A, I, and R. We have stopped talking about FAIR and removed it from our replacement H2020 DMP template. It’s so confusing”

“I support the idea of FAIR, but have reservations about the amount of details which are asked. If the added value of making data FAIR becomes more clear, people will automatically follow best practices in their domain.” Also, it is recommended that researchers who practise Open Data should be adequately compensated or rewarded for the time spent on this task, particularly early-career researchers.”

One respondent, moreover, commented as follows:

“Our institution is a technical university and therefore innovation is a crucial part of the research performed here. Often, researchers and even support teams from the innovation office see the ORD pilot as a threat of innovation, specifically patenting. We have shared the message that the pilot does not intend to jeopardize the innovation and patent ability of the research output, and that embargo periods for publication of data should protect that. But, it would be useful to get information/education/webinars/strategy for researchers and/or innovation officers from the EC on how to profit from good practice in RDM, but also open access to data. Where are the challenges and how to overcome them, etc.

Appendix I - Infographic of the survey results



Results of the

SURVEY ON THE HORIZON 2020 DMP TEMPLATE

289 RESPONDENTS



- DMP writers 50%
- DMP support staff 60%
- Both 10%

overall experience



- of 189 respondents
- positive 60%
 - not applicable 24%
 - negative 16%



understand the
FAIR concept

yet practical implementation remains difficult

almost half

would openly
publish a DMP



Yes if...

And even more would
do so if certain
conditions were met
such as confidentiality.



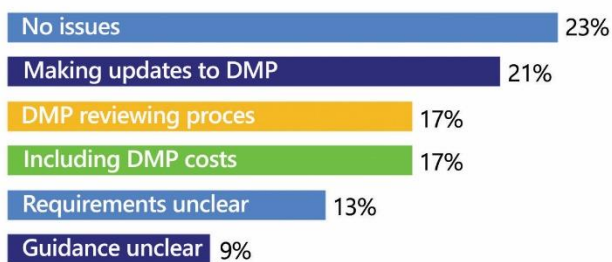
Top 5 priorities for a DMP tool:

1. Suggest relevant **standards** for my field and data type
2. **Drop-down options** based on good practice per discipline
3. Give more **examples** or suggested answers
4. Include **discipline specific** guidance and tailoring
5. Recommend **repositories or tools** that I can use



"As **open data** is a crucial issue
in recent science policy,
the compilation of a DMP helped me to become
familiar with the **respective requirements**."

Issues encountered when following H2020 guidelines



Our recommendations for H2020 DMPs:

- Revise the DMP template structure
- Reduce technical terminology
- Provide discipline-specific guidance
- Offer example DMPs and costings
- Clarify DMP review processes



Appendix II - Recommendations and next steps

There were various suggestions or requests for further support and clarification emerging from the survey responses. We have grouped these into key areas of activity and made suggestions as to how they could be addressed. In responding to these changes, we recommend that the European Commission draws in support from projects such as OpenAIRE and OpenAIRE Advance, EOSC-Hub and FOSTER Plus, as well as the ESFRIs and the Fair Data Expert Group, which has been tasked with providing advice on Data Management Plans.

1. Clarify EC requirements for DMPs

Information on the Open Data Pilot and the requirements for DMPs is held in different documents. Some of the most useful guidance (including a definition of data and requirements of the pilot) is actually contained in the “Open access to publications and research data in Horizon 2020” document. Although this document is linked to early in the “Guidelines for FAIR Data Management” document, it is very likely that researchers will miss one or more of the documents as they already have to deal with so much documentation. Ideally all the information provided on the data pilot and DMPs would be in a single overview to remove any inconsistencies or differences in emphasis. Formal announcements should be made on data management listservs (e.g. research-dataman@jiscmail.ac.uk) and specific groups alerted (e.g. OpenAIRE NOADs and DMP tool providers) whenever these guidelines are updated.

Certain aspects of the requirements were also found to cause confusion. For example, questions were raised as to whether the set of questions in the ‘Template’ or the ‘Summary Table’ part of the FAIR Data Management guidelines should be used. Each is very similar but has slightly different questions. Several respondents commented that the DMP wasn’t relevant for their work, so more examples should be given on when they aren’t required, for example on certain programmes or if no data are created. Advice and training on this should also be provided for project officers (see section 7 on review)

- Collate all data-related and DMP guidelines in a single document
- Provide public announcements and alerts to key groups when guidelines change
- Be more explicit about when a DMP is and isn’t required
- Clarify exactly what set of questions should be answered in DMPs
- Provide a docx or rtf template to use, as well as supporting online tools

2. Revise the DMP template structure

A lot of detailed feedback was provided on the DMP template, highlighting issues in terms of repetitive or redundant questions, terminology and confusing concepts. We have suggested two changes to address these: first to revise the structure of the DMP template to regroup and order questions in a more logical way that avoids repetition and ideally allows users to skip irrelevant sections; and second to improve the content and terminology to ease understanding.

Many of the issues related to the template being structured according to the FAIR principles. While this is a useful concept at a high-level, ordering questions in this way causes overlap and redundancy as certain concepts like metadata cross-cut the issues of ‘findability’ ‘interoperability’ and ‘reuse’. We propose that questions could be more usefully grouped thematically or by key activities (e.g. collecting or generating data, developing and documenting metadata, and data sharing).

- Regroup questions according to key activities and order in a more logical fashion

- Identify which questions should be answered when, highlighting a subset of primary questions that should be addressed at the six month point
- Where possible, indicate secondary questions that can be skipped if not relevant e.g. if the researcher has identified the data are not sensitive, then questions around data access committees can be ignored
- Include more yes/no questions, dropdown options or API integrations querying external databases to avoid so many freetext responses
- Identify questions that will support evaluation and formulate them in a structured way to enable automated compliance checks in future

3. Simplify the DMP content and terminology

Survey respondents identified over 40 terms that caused confusion, with 'interoperability', 'ontologies', and 'metadata' being chief amongst them. At minimum a glossary of terms should be provided that offers definitions and examples to contextualise these terms. However where possible, the language used should be simplified to reduce technical jargon.

The coverage of the DMP template is also very comprehensive, with over 40 questions addressing topics as diverse as the origin of the data, how it will be named and versioned and how the identity of the person accessing the data will be ascertained. Some questions are very broad while others are specific and only relevant to a subset of projects. Questions could usefully be graded with a higher-level, more generally applicable question asked to all, while more specific questions arise as and when relevant. For example all projects should state which repository the data will be deposited in, but questions around data access committees and identifying the person accessing the data are only relevant when it is available under restrictions or via secure data services, rather than as open data.

- Simplify the terminology used where possible
- Provide a glossary of terms to assist researchers in understanding concepts
- Offer example answers (see point 4), particularly for questions with more technical terms so these can be contextualised
- Seek to shorten the number of questions being asked
- Provide some hierarchy or routing to the structuring of the questions so more niche concerns are only asked when relevant to the project context

4. Provide discipline-specific guidelines and example answers

One of the most frequent requests for support was discipline-specific guidelines. Projects would really welcome example answers or ranges of options based on good practice for their field. Existing work in this area such as the Science Europe domain protocols could be built on, and collaboration with domain data centres and discipline-specific groups such as learned societies should be pursued.

- Provide discipline-specific guidelines
- Offer example answers based on good practice in each domain
- Give dropdown options or ranges pertinent to each field

5. Encourage the publishing of DMPs and collate examples

Survey respondents demonstrated a willingness to publish DMPs, and indeed many projects have already made them available via their own websites, publishers such as RIOjournal or repository platforms like Zenodo. The EC should promote the notion of making DMPs public and of preserving the final version along with the data. Early publishing would allow information to be reused to increase data discovery and improve resourcing of RDM support. A registry service should be created to record and collate these DMPs. In addition, a small subset of 'approved' DMPs should be collated to give projects a sense of what represents a 'good' DMP for the various

research areas supported by the EC. Collaborating with external reviewers and EC infrastructure projects will help to identify 'good' examples.

- Encourage projects to publish their DMPs
- Offer a DMP registry service
- Provide a library of 'approved' DMPs representing the range of funded research

6. Facilitate the inclusion of RDM costs in grant applications

The current approach to DMPs is at odds with the need to consider RDM costs and write these in at the proposal stage. The EC should prompt projects to consider RDM costs as part of the budget at the proposal stage. This need not mean drafting a DMP at the proposal stage, but certain key questions that have an impact on the data resourcing requirements (e.g. scale of data, ethical concerns, data security etc) should be flagged to ensure associated costs are considered and incorporated.

Researchers are often reluctant to include RDM costs for fear their proposal will seem uncompetitive. Many questions were also raised about exactly what is an eligible cost and how to calculate figures for these. Worked examples are urgently needed to clarify the range of activities that could be included and to demonstrate how calculations can be made as to the appropriate level of resourcing needed. Guidance from the Wellcome Trust and LCRDM provides useful models here. It is also critical to brief reviewers so they question why RDM costs aren't included when data management is a significant activity for the project. This will help promote the culture change required to include RDM costs as a matter of course.

- Raise a few RDM questions / alerts at the grant proposal stage to prompt researchers to consider and include associated costs
- Provide worked examples that demonstrate what to include as RDM costs and how
- Brief reviewers on eligible costs and provide guidelines so they raise concerns if these aren't included

7. Improve DMP review practices and share guidelines

The topic of review came up throughout the survey and respondents sought a lot of clarifications on exactly how the EC is reviewing DMPs. Although the process varies as this is still being piloted, we recommend formalising the internal guidelines and endorsing the existing assessment framework so it becomes a common approach across units. Existing training provided by the FOSTER Plus project, DCC and external reviewers should be repeated so all project officers have access to support. The FOSTER Plus webinar for project officers planned for January 2018 and the proposed internal guidelines and FAQs will be key resources to embed common practices across the Commission.

It would be helpful to publish a statement on how DMPs are being reviewed, and ideally release the assessment framework. Since the framework is based on the template, and changes to this will need to be taken into account. At recent FOSTER Plus organised training courses for EC project officers held in November 2017 we discussed highlighting a subset of criteria that should be more heavily weighted in the review process - i.e. questions which if not addressed or dealt with poorly, have a far greater impact on the management and sharing of data. This would include aspects like the chosen repository, approach to metadata, data licence and use of standards. This should be done in parallel with adjustments to the template structure proposed in recommendation 1.

- Endorse the assessment framework to ensure consistent use across all units and highlight a subset of primary questions within it to aid review
- Circulate guidelines on how to review DMPs together with the assessment framework and FAQs to advise on practical implementation of the ORD pilot

- Continue to provide internal training on reviewing DMPs in collaboration with FOSTER Plus and other relevant groups
- Make a public statement on how DMPs are being reviewed, ideally releasing the assessment framework at the same time



Appendix III - List of survey questions

Q1. For the purpose of the survey we distinguish two roles. If you belong to both groups, please check the appropriate box and use the open text fields for extra information when your experience in these roles varies. [multiple choice]

- “DMP writer”: researchers, principal investigators, project coordinators and others who write and submit a DMP to the Commission.
- “DMP support staff”: research support staff who provide expertise and practical support to DMP writers; this includes the reviewing of draft DMPs.

Q2. Please note your name and affiliation. We do not publish combinations of names and individual answers.

Q3. Did you find the process of writing / supporting a Horizon 2020 DMP positive or negative? [positive, negative, or free text]

- Positive
- Negative
- Not applicable

Q4. To what extent do the following statements represent your experience of using the H2020 template? [5-point scale per statement]

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
It is a very useful template.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The structure is clear and it groups questions into helpful categories.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I like the level of guidance and contextual information that is provided.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
There are too many free text answers. I need more dropdown options.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
There are too many questions!	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Many questions are irrelevant. I'm not sure why they are being asked.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I don't understand what FAIR means.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q5. Terminology - please note any terms that are unknown, unclear or confusing [free text]

Q6. Questions - please note any questions that are confusing, inconsistent, or redundant [free text]

Q7. Guidance - please note any guidance that is missing or could be improved [free text]

Q8. Coverage - please note any topics that are missing or ones that should be removed [free text]

Q9. Process - please note any issues encountered following the EC guidelines e.g. knowing when a DMP is due, how it will be reviewed, how to include costs etc [free text]

Q10. Other - any other issues or suggestions regarding the template you would like to make [free text]

Q11. In an ideal world, what would you like a DMP template or associated tool to do for you? Please rank your top 5 priorities below (one response per column) [select and rank 5 out of 10 options, plus free text]

	Top priority	2nd priority	3rd priority	4th priority	5th priority
Suggest relevant standards for my field and data type	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Provide dropdown options based on good practice for my discipline	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Give more example or suggested answers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Include disciplinary guidance and tailoring	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Recommend repositories or tools that I can use	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Alert repositories that I plan to deposit there	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Share information with my university or data services to plan storage and support	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Connect to other research tools and exchange data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pre-fill and automatically update information	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Publish information about my data to help others discover and reuse it	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Other (please specify)	<input type="text"/>				

Q12. Would you openly publish your DMP [yes, no, or free text]

Q13. The European Commission plans to add disciplinary guidance to its generic H2020 DMP template. What standards, tools, databases, repositories or other support and best practices should be referenced for your discipline? [free text]

Q14. What are the main suggestions you would give to the European Commission about its approach to DMPs in the Open Data pilot? [free text]

Q15. Have you given feedback to (other) DMP writers on their draft DMP? (If not, you will automatically go to the last questions of this survey, after clicking on the 'Next'-button/) [yes or no]

Q16. What kind of feedback did you give? Please tick all that apply. [multiple choice plus free text]

- You explained the *rationale behind one or more questions* in the template, because you noticed/assumed they were misunderstood
- You explained one or more *terms* in the template, because you noticed/assumed they were misunderstood
- You referred to more *guidance* about one or more topics in the template
- You asked for more information on one or more topics, because information was lacking or you considered it insufficient – with regard to the version of the DMP (initial, mid-term, final)
- You remarked on one or more *inconsistencies* within the DMP
- Other (please specify)

Q17. How did the process of reviewing and supporting researchers with a H2020 DMP differ from assisting them to write plans for other funders or contexts? [single choice including free text]

- There was no difference
- The H2020 process was quicker and easier
- The H2020 process took much longer and was more complicated
- I was confused by the H2020 template and felt less able to advise
- N/A - I have only supported H2020 DMPs
- Other (please specify)

Q18. How much time do you need to review a draft Horizon2020 DMP on average? [single choice]

- Less than 30 minutes
- Between 30-60 minutes
- Between 60-90 minutes
- More than 90 minutes
- I don't know / I don't often review DMPs

Q19. Do you have any other recommendations about the Guidelines for FAIR Data Management? [free text]



Appendix IV - Request to circulate the survey

Dear colleagues,

The OpenAIRE project is collecting feedback on the Horizon 2020 template for Data Management Plans. Please complete the survey at: https://www.surveymonkey.com/r/OpenAIRE_DMP_survey

Since this is an open data pilot, it's your opportunity to respond and help shape future policy. We're really keen to hear your feedback on the Commission's approach to DMPs and the FAIR DMP template in particular. Please give us specific feedback to help us understand what works well and what doesn't so recommendations can be put back to the European Commission.

OpenAIRE will publish a summary of the responses to the survey and recommendations on the project website. The survey results will also be used as an input to the [Fair Data Expert Group](#), which will contribute to an evaluation of the Horizon 2020 DMP template and suggest future revisions.

We'd like to hear from

- researchers, principal investigators, project coordinators and others who have written Horizon 2020 DMPs, as well as from
- research support staff who provide expertise and practical support to DMP writers.

Many thanks for your participation. Thanks also to the Dutch National Coordination Point RDM for the questions we borrowed from their recent survey on DMP reviewing experiences.

Our [survey](#) will close on June 21st 2017. Please respond before then and help us by circulating it to relevant lists.

Kind regards,

Research Data Management team in OpenAIRE