Assessing the Value and Usage of Data Management Planning and Data Management Plans Within the U.S. Geological Survey

Assessing the Value and Usage of Data Management Planning and Data Management Plans Within the U.S. Geological Survey

By Madison L. Langseth, Elizabeth A. Sellers, Grace C. Donovan, and Amanda N. Liford

Science Synthesis, Analysis, and Research Program

Open-File Report 2023–1069

U.S. Department of the Interior
U.S. Geological Survey

U.S. Geological Survey, Reston, Virginia: 2023

Contents

Figures

Tables

Abbreviations

BAO	Bureau Approving Official
DMP	data management plan
DOI	Department of the Interior
FAIR	findable, accessible, interoperable, and reusable
FSP	Fundamental Science Practices
IT	information technology
maDMP	machine-actionable data management plan
N	number of responses
PC	Program Coordinator
PD	position description
Q	question
RDM	research data management
SM	Survey Manual
USGS	U.S. Geological Survey

Assessing the Value and Usage of Data Management Planning and Data Management Plans Within the U.S. Geological Survey

By Madison L. Langseth, Elizabeth A. Sellers, Grace C. Donovan, and Amanda N. Liford

Abstract

As of 2016, the U.S. Geological Survey (USGS) Fundamental Science Practices require data management plans (DMPs) for all USGS and USGS-funded research. The USGS Science Data Management Branch of the Science Analytics and Synthesis Program has been working to help the USGS (Bureau) meet this requirement. However, USGS researchers still encounter common data management-related challenges that may be reduced or eliminated by better planning. In 2021, USGS staff were given a series of surveys aimed to better understand current data management planning practices, perceptions, and needs. The survey results indicated that adoption and integration of data management planning and DMPs into USGS research project workflows are broad, if inconsistent, across USGS Science Centers and programs. The USGS Science Data Management Branch can help improve clarity and guidance on the purpose, intended audience, content, workflows, and evaluation processes for DMPs. It would also be beneficial to provide additional supporting cyberinfrastructure to support DMP activities. Survey responses indicated it would be beneficial for the Science Data Management Branch to develop a strategy, other than through DMPs, for teaching and encouraging good data management practices. Although these surveys were an opportunity for USGS staff to provide feedback on their experiences, the surveys may also have revealed the desire for more frequent evaluations, cross-disciplinary communication, and training on research data management and DMP development and integration, in the context of USGS policy, Fundamental Science Practices requirements, and overall Bureau expectations. Data management-related roles such as data manager or steward, information technologist, and repository manager may need to be formally recognized as skilled professional career positions within the Bureau. At a minimum, the best practice for USGS would be to create and maintain DMPs as living documents, integrated with existing systems that are broadly accessible to all stakeholders, and include quantitatively measurable benefits tied directly to a clearly defined purpose.

Introduction

The Science Data Management Branch within the Science Analytics and Synthesis Program of the U.S. Geological Survey (USGS) Core Science Systems Mission Area provides Bureau-wide leadership to optimize and share USGS science data management practices and workflows for describing, preserving, and publishing USGS science data and related information products. The branch takes the following four-step approach to achieve this mission:

1. developing enterprise tools such as ScienceBase (https://sciencebase.usgs.gov), a trusted digital repository (USGS, 2016; Hutchison and others, 2021) for the Bureau and USGS Science Data Catalog (https://data.usgs.gov), a metadata repository for published data from the Bureau;

2. building communities of practice to bring together scientists, data and metadata managers, and technology developers from across the USGS to share ideas, learn new skills, and establish best practices for data management and data integration;

3. contributing to the development and promotion of best practices in data management through educational webinars, training, and resources such as the USGS Data Management website (https://www.usgs.gov/data-management); and

4. participating in Department of the Interior (DOI) and USGS data management policy development activities and working to closely align our work with that of the Bureau, DOI, and other Federal government requirements that support making USGS data findable, accessible, interoperable, and reusable (FAIR; Wilkinson and others, 2016).

In an effort to advance USGS data management planning practices and workflows, the Science Data Management Branch distributed surveys designed to assess the current

levels (as of March 2021) of participation, implementation, challenges, and opportunities for improvement in relation to the creation and use of data management plans (DMPs).

Background

On February 22, 2013, the Office of Science and Technology Policy issued a memorandum directing Federal agencies "to develop a plan to support increased public access to the results of research funded by the Federal Government" (Office of Science and Technology Policy, 2013). In response, the USGS developed a Public Access Plan that requires DMPs for all USGS and USGS-funded research (USGS, 2016). In February 2015, the USGS also issued "Instructional Memorandum OSQI–2015–01, Review and Approval of Scientific Data for Release" (USGS, 2015), which documented a series of data management-related policies. In January 2017, the Instructional Memorandum was released as USGS Survey Manual (SM) chapter "SM 502.6—Fundamental Science Practices: Scientific Data Management" (USGS, 2017a). The SM chapter instructs USGS staff to include a DMP as part of their project work plan (USGS, 2011) prior to initiation of the project. Approved USGS DMPs must outline methods for managing, releasing (for example, in appropriate long-term repositories using nonproprietary, open formats), and preserving digital data (USGS, 2016). Although SM 502.6 (USGS, 2017a) and the Public Access Plan (USGS, 2016) describe what should be included in a DMP, they do not communicate why DMPs are required.

The USGS Science Data Management Branch helps the Bureau meet the Fundamental Science Practices (FSP) requirements detailed in SM 502.6 by providing guidance to scientists on the USGS Data Management website and data management tool recommendations such as the USGS Data Management Plan Editor/Manager (DMPEditor). Despite having DMP policies and general guidance on the Data Management website, USGS researchers still encounter challenges at the point of data release that might be reduced or eliminated by better planning at the beginning of research projects. Some USGS Centers have fully embraced the DMP concept, whereas others are still determining how to best integrate DMPs into their project planning activities. Anecdotal information acquired informally through the USGS Community for Data Integration's Data Management Working Group revealed a glimpse of the challenges researchers face and their needs in relation to DMP development and use, but our understanding of these challenges across the Bureau is incomplete.

Therefore, in the spring of 2021, the Science Data Management Branch distributed surveys to collect baseline data about data management planning practices, perceptions, and needs from USGS staff across the Bureau in a variety of roles. These baseline data enable us to understand the state of data management planning within the USGS and suggest potential ways to refine the implementation and increase the value of DMPs.

Methods

The following sections describe the survey development from design and testing to distribution and analysis. The Science Data Management Branch also interviewed survey participants as a followup to the surveys, and the interviewee selection process is described.

Survey Design

Four surveys (see appendixes 1–4, available at https://pubs.usgs.gov/publication/ofr20231069) were designed to collect information on the current level of participation (as of March 2021), implementation, challenges, and opportunities for improvement in relation to data management planning and the development and curation of DMPs. The surveys were divided into three main sections. The first section consisted of questions related to project documentation and processing through proposals, work plans, DMPs, approvals and reviews, and project tracking systems. The second section consisted of questions specifically about DMPs, their usefulness, and related challenges. The third section focused on questions about demographics, affiliations, and awareness of DMPs and related project planning topics. Each survey provided an opportunity for the participants to share additional feedback by volunteering to be interviewed on the topics addressed in the survey and by writing their feedback in a freeform text field. Microsoft Forms was used as the survey development and deployment tool, and a Power Automate workflow was used to export the response data from each survey to a Microsoft Excel workbook.

Survey Groups

To cater the survey questions to the participant's role, survey participants were categorized into four distinct groups: researchers, data managers and information technologists, Center Directors, and Program Coordinators and Bureau Approving Officials. Survey questions were rephrased for each group but tailored to collect the same type of information regarding data management planning and DMPs. The distribution lists for the four survey groups were determined as follows:

Researchers (1,484 individuals):

a. USGS Research Grade Evaluation scientists;

b. scientists that have contacted the USGS ScienceBase Data Release Team by email between 2018 and 2021; and

c. members of the Community for Data Integration Data Management Working Group who were not already identified as belonging to the data managers and information technologists survey group (see "Data managers and information technologists" breakdown).

Data managers and information technologists (hereafter discussed as "data managers") (529 individuals):

  a. Science Center staff listed as points of contact associated with USGS data releases hosted in a USGS trusted digital repository (for example, ScienceBase);

  b. members of the USGS Water Mission Area Data Manager Forum;

  c. members of the Ecosystems Mission Area Data Managers Team; and

  d. members of the USGS information technology (IT) email lists, which includes all USGS IT personnel.

Center Directors (92 individuals):

  a. Directors and Acting Directors of USGS Science Centers.

Program Coordinators (PCs) and Bureau Approving Officials (BAOs) (26 individuals):

  a. USGS staff with PC and BAO roles as maintained by USGS human resources.

All survey group distribution lists were cross-referenced to ensure that individuals would receive only a single email invitation to respond to the survey most appropriate to their role. Subsequent recommendations from survey participants to invite additional individuals to respond to a survey were also collected. A total of 2,131 individuals were invited to take the surveys.

Survey Pilot Testing

Each survey was pilot tested by at least one representative from each survey group. Invitations were sent to survey pilot testers, and their survey responses and feedback were received February 21–23, 2021. The information gathered during this pilot testing phase was used to clarify and validate the surveys' structure, formatting of questions, and response options.

Survey Distribution

On February 24, 2021, the Associate Director of the USGS Core Science Systems Mission Area informed the USGS Executive Leadership Team of the intent to distribute the surveys to USGS staff. Formal written permission from the Executive Leadership Team to distribute the surveys was received on March 1, 2021. Microsoft Word's Mail Merge function was used to create and send standardized individual email notifications to each recipient. Invitations to complete the surveys were emailed to each of the survey groups on March 2 and 3, 2021. On March 17, 2021, reminder emails were sent to those who had not yet completed the survey. The survey completion deadline was March 26, 2021.

Personally Identifiable Information

Where free-text responses included information that could potentially reveal a respondent's identity, affiliation(s), or any other information that might affect their anonymity, the information was manually replaced by USGS Science Data Management staff with a generic phrase (for example, [Center name], [program name], [system name], [person name]). Sensitive information (for example, first and last names, email addresses, and Science Centers) was also redacted from the supporting data as part of the processing and preparation for formal data release (Langseth and others, 2023).

Data Analyses

Jupyter Notebooks (Langseth and others, 2023) were used to perform data manipulation and analyses, basic summary statistics on survey responses, and Pearson's chi-square tests of independence (χ^2) to examine the relations between factors associated with DMP creation, maintenance, and usefulness. In all surveys, the response "I don't know" was provided as an additional option to Yes or No questions. For analysis of responses to this question type, particularly in the case of the researcher survey, "No" and "I don't know" responses were treated as a single category. Additionally, the lack of a resource and the lack of knowledge about a resource's availability were considered to have the same effect.

Interviewee Selection

After the survey closed, interviews were performed with willing participants to follow up on notable themes emerging from the researchers and data managers surveys. Each of the surveys gave participants the opportunity to indicate their willingness to be interviewed (appendixes 1 and 2, question [Q] 29). Interviews were initiated to learn more about the following: (1) what data management resources and training were used by researchers and their preferred formats for them; (2) how and why researchers are using DMPs as living documents and their processes for updating them, if applicable; and (3) the perception of and culture around DMPs as experienced by researchers and data managers (appendixes 5 and 6, available at https://pubs.usgs.gov/publication/ofr20231069).

Interview participants were selected based on their responses to certain survey questions. Ten data managers were interviewed based on their response to the question "Have you found project DMPs to be useful to your work?"; five who responded that DMPs are very or somewhat useful and five who responded that DMPs are not useful at all, somewhat not useful, or neutral. Researchers from each of the categories described in table 1 were interviewed for a total of 10 researcher interviews.

Table 1.   Criteria for selected researcher interviewees.

[Fields marked as not applicable (NA) indicate there were no survey participants who met the response criteria. Questions in the column headings are questions 15, 9, and 11 in appendix 1. DMP, data management plan]

Have you found project DMPs to be useful to your projects?	How are the DMPs maintained?	Does your Center have a documented process for creating project DMPs?	Number of interviewees selected
DMPs are very or somewhat useful	As static documents	Yes	NA
		No	3
	As living documents	Yes	2
		No	1
DMPs are not useful at all, somewhat not useful, or neutral	As static documents	Yes	NA
		No	4
	As living documents	Yes	NA
		No	NA

Results

The researchers survey (appendix 1) received 486 responses, a response rate of 32.6 percent. The data managers survey (appendix 2) received 159 responses, a response rate of 30.1 percent. A large percentage (44.8 percent) of survey respondents in the data managers survey group indicated that they did not have a role related to DMPs. Consequently, their responses to subsequent questions in the survey were "I don't know," indicating that the USGS IT All email listserv was likely too broad of a group to include in this survey. Therefore, in the remaining sections, we only report results from the 89 data manager survey group respondents who indicated that they do have a role related to DMPs. The Center Directors survey (appendix 3) received 34 responses, a response rate of 32.6 percent. The PCs and BAOs survey (appendix 4) received eight PC responses, a response rate of 53.3 percent, and nine BAO responses, a response rate of 75.0 percent.

BAOs and PCs survey group responses constituted a sample size that was insufficient to support quantitative analysis. However, a qualitative review of their responses showed that BAOs and PCs are less involved in DMPs in their positions, and many of their responses cited their perceptions or experiences from their perspective while in previous roles as researchers or Science Center Directors. In this respect, their responses further confirmed the perceptions seen in the responses from the other survey groups.

Three individuals responded to both the researcher and data manager surveys and one individual responded to both the researcher and Center Director surveys because they had experience in both roles and could provide insight from both perspectives. Thus, their responses were retained in analyses of both surveys.

Overall Demographics

The distribution of respondents across all survey groups corresponded roughly to the distribution of USGS staff across DOI Active Directory departments, which equate to USGS Regions or Mission Areas (fig. 1). The DOI Active Directory is a system used to manage USGS personnel information.

The initial email distribution lists consisted of representatives from 107 USGS Science Centers. Researcher respondents represented 75 Science Centers. Data manager respondents represented 62 Science Centers, and Center Director respondents represented 33 Science Centers. A total of 80 out of 114 Science Centers were represented in the responses received for all three surveys. The eight PCs that responded to the survey represented the USGS Ecosystems (four participants), Natural Hazards (one participant), and Water Resources (three participants) Mission Areas.

Roles Related to Data Management Plans

Data managers and Center Directors were asked what their roles were with respect to DMPs at their Centers (appendix 2, Q13, and appendix 3, Q20). The question was multiple choice for both groups, but data managers were also able to enter free-text responses to this question. Center Directors were not given the free-text option because their role is already defined by USGS FSP (USGS, 2017b).

Data managers who reported having roles related to DMPs represented 50 USGS Science Centers (43.9 percent of all Science Centers). Of the data manager respondents who considered themselves to have some sort of role related to DMPs at their Centers, the highest percentage (38.2 percent) reported that they "Ensure project DMPs are created" and
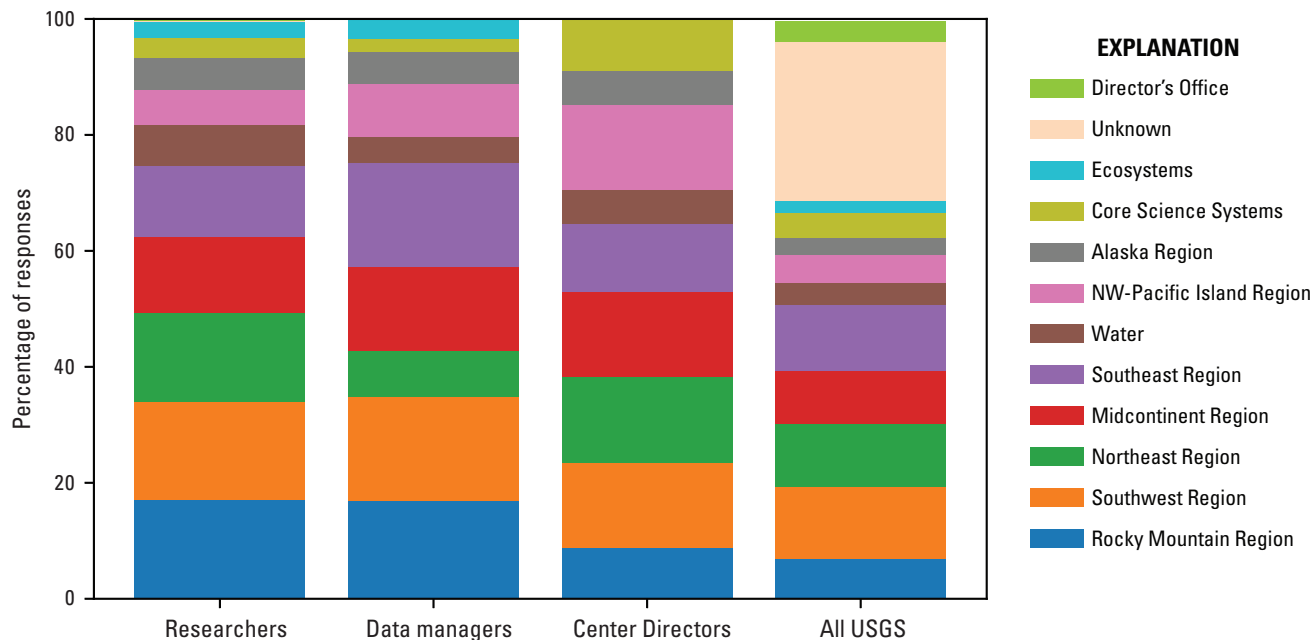
Figure 1. Distribution of researchers, data managers, Center Directors, and all U.S. Geological Survey personnel across Department of the Interior Active Directory departments. Active Directory uses "UNKNOWN" for a person's department when that information is not provided within a given person's record. NW, northwest; USGS, U.S. Geological Survey.

"Review or approve project DMPs" (fig. 2). Free-text answers from data managers primarily included more detailed information on their specific roles related to DMPs. For example,

"I manage the end point for the data distribution."
USGS data manager

"[I] answer the questions on how to create DMPs and enter information into our internal system."
USGS data manager

Data managers who reported having some role related to DMPs represented 37 different DOI Active Directory position titles (table 2), and the most common was "IT Specialist." A minority (ten individuals) had position titles that included the word "data," indicating that data managers are likely taking on this DMP role as collateral duty.
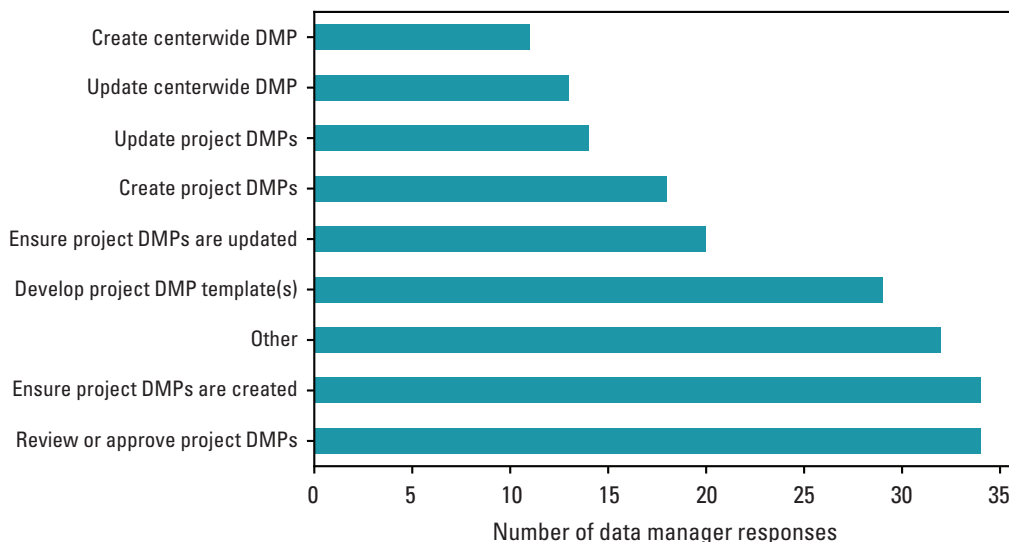
Figure 2. Distribution of U.S. Geological Survey data manager roles in relation to data management plans (DMPs). Number of respondents was equal to 89 (Appendix 2, question 13).

Table 2.   Frequency of Department of the Interior Active Directory position titles represented among data managers that report having a role related to data management plans.

[Position titles that include the word "data" are highlighted and represent the minority of data managers' position titles. IT, information technology; mgmt, management; datamgt, data management; custspt, customer support]

Active Directory position title	Frequency
IT specialist	14
Hydrologist	12
Hydrologic technician	8
Geographer	5
Supervisory IT specialist	4
Data management specialist	4
Supervisory hydrologist	4
Physical scientist	3
Cartographer	2
Computer scientist	2
Geologist	2
Geophysicist	2
IT specialist (sysadmin)	2
Physical science technician	2
Administrative officer	1
Biologist	1
Branch chief	1
Cartographic technician	1
Data and IT coordinator	1
Data management, IT specialist	1
Data modeler	1
Data scientist	1
Ecologist	1
Geospatial analyst	1
Hydrologist (acting reports specialist)	1
Information education specialist	1
IT specialist (data mgmt)	1
IT specialist (datamgt/custspt)	1
Management analyst	1
Program analyst	1
Project manager	1
Supervisory biologist	1
Supervisory wildlife biologist	1
Surface water specialist	1
Technical information specialist	1
Unknown	1
Wildlife biologist	1

According to USGS SM 502.6, Center Directors or their designees are responsible for "ensuring the development of data management plans for all new research proposals and the updating of these plans." When asked about their role(s) with respect to DMPs at their Center, most Center Directors (61.3 percent) responded that they delegate this DMP authority (see USGS, 2017b) to another person or team (fig. 3).

Creation of Data Management Plans

When responding to the first section of the questionnaire, researchers were asked to think about one of their most recent projects. A little over one-half of researchers (51.9 percent) reported that a DMP had been created for the project, either by the respondent (31.9 percent) or by another person (20.0 percent). Fewer respondents (8.8 percent) said that their project work plan referenced an existing DMP. Finally, 39.3 percent of respondents reported that no DMP existed for their project.

In subsequent analyses, responses were grouped into two categories: "DMP created" (including existing DMPs) and "No DMP created." These two categories were also used for branching within the survey (fig. 4). For example, respondents in the "DMP created" category were asked several followup questions related to access and maintenance of the DMPs (appendix 1, Q8–10).

Data managers were asked if projects at their Centers generally have DMPs (appendix 2, Q7). Data manager respondents predominantly reported that "Most" (31.5 percent) or "A few" (31.5 percent) projects at their Center have DMPs (fig. 5).

Center Directors were asked whether they require projects at their Centers to have DMPs (appendix 3, Q16). Most Center Directors (64.7 percent) answered that "Yes, all of them (projects)" are required to have a DMP (fig. 6).

For those Center Directors who responded that "Most" or "A few" of their Center's projects have DMPs, a followup question was asked to determine how they decide which projects need a DMP. Some free-text responses to this followup question indicated that the primary deciding factor on whether projects have DMPs is based on the funding source and whether the funder requires a DMP, whereas others stated that the type of project affects whether there is a DMP. For example, fieldwork-based projects will have a DMP, but laboratory-based projects may not, or projects funded by different USGS Mission Areas may have different levels of requirements. Other responses indicated that all projects should have DMPs, but a lack of resources or Center guidance causes them to fall short of this requirement.
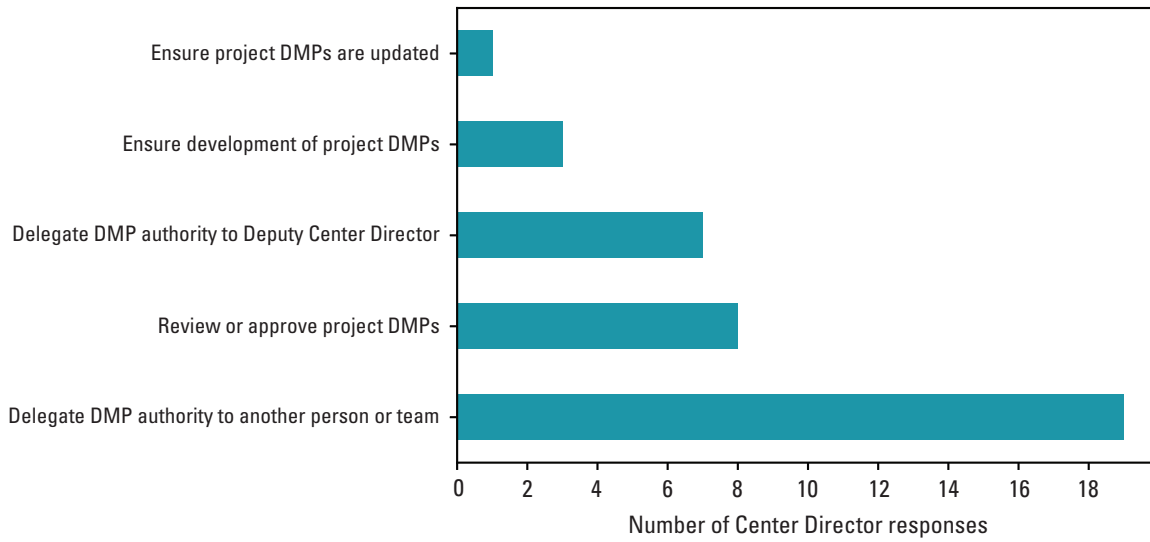
Figure 3. Roles reported by Center Directors in relation to data management plans (DMPs). Number of respondents was equal to 31 (Appendix 3, question 20).
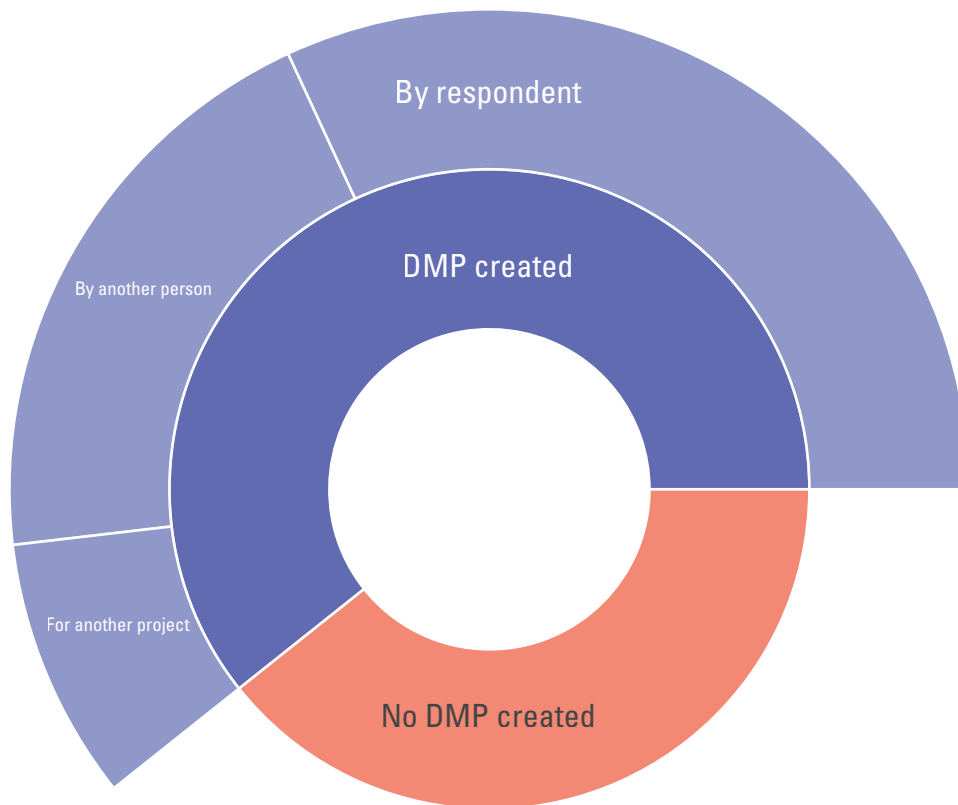
Figure 4. Figure illustrating how responses related to data management plan (DMP) creation were grouped for subsequent analyses (Appendix 1, questions 8–10).
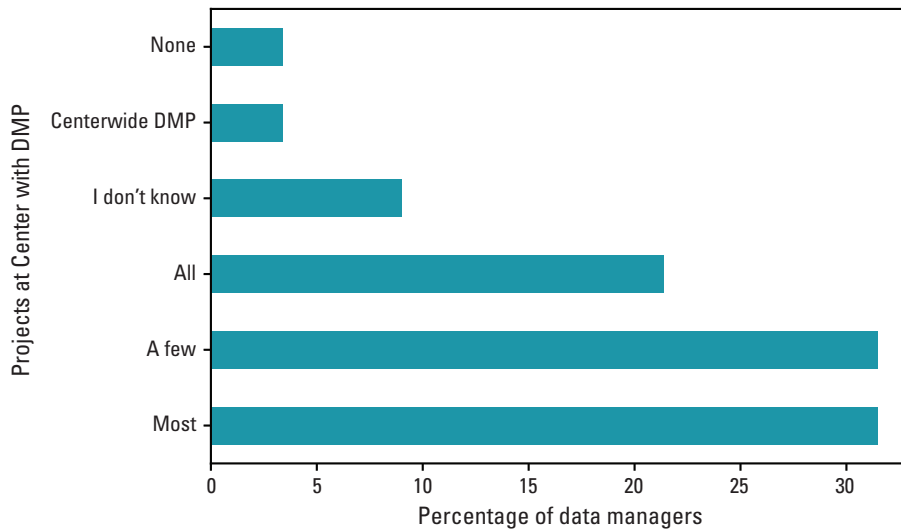
Figure 5. Prevalence of projects with data management plans (DMPs) at data managers' Centers. Number of responses was equal to 89 (Appendix 2, question 7).
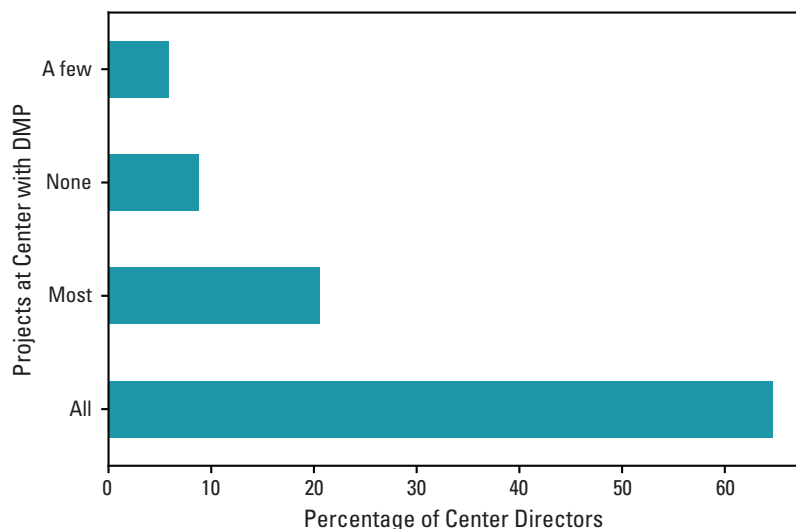
Figure 6. Prevalence of projects required to have a data management plan (DMP) at Center Directors' Science Centers. Number of responses was equal to 34 (Appendix 3, question 16).

Center Data Management Plan Processes, Policies, and Resources

In each survey, respondents were asked several questions about their awareness of DMP processes and resources at their Center. One of the first questions was about whether their Science Center has a documented process for creating DMPs. Thirty-six of the 49 Science Centers represented in the data manager responses reported their Center does have a documented process for creating DMPs. Twenty of the 34 Science Centers (approximately 58 percent) represented in the Center Director responses reported their Center does have a documented process for creating DMPs. When considered along with responses from researchers, 64 out of the 80 Science Centers (80 percent) have at least one person reporting their Center has a documented process for creating DMPs. However, responses from within one-fourth of Centers (20) represented in survey responses indicated there is some disagreement among respondents from the same Center regarding whether documented processes exist. For example, in some cases, a researcher may have reported a positive response, but a data manager or Center Director from the same Center may have reported a negative response regarding whether there is a documented process for creating DMPs at their Center.

One-half of researcher respondents (50.5 percent) indicated their Center does have a documented process for creating DMPs, whereas 41.2 percent of researchers were "not sure" and 8.2 percent answered "No" (their Center does not have a documented process for creating DMPs; fig. 7).

Data Management Plan Creation Support

Less than one-half of researchers (48.2 percent) indicated someone is available at their Center to help them create a DMP (fig. 8). Slightly fewer researchers (37.9 percent) were not sure if help is available, and 13.8 percent reported that help is not available. Researchers who reported having help represent 61 unique Science Centers across the USGS. Data

managers and Center Directors were not asked about the availability of someone to help create DMPs because data managers often fulfill this role for their Center and Center Directors are not expected to be involved in DMP creation.

Data Management Plan Update Policy

When asked if their Center has a documented policy (versus a process or procedures) for updating DMPs, most researchers did not know (65.6 percent), less than one-fourth of respondents (22.1 percent) answered "Yes," and 12.4 percent answered "No" (fig. 9). Researchers who reported having a documented policy for updating DMPs represented 47 unique USGS Science Centers. Data managers and Center

EXPLANATION

Yes (50.5 percent)

I don't know (41.2 percent)

No (8.2 percent)
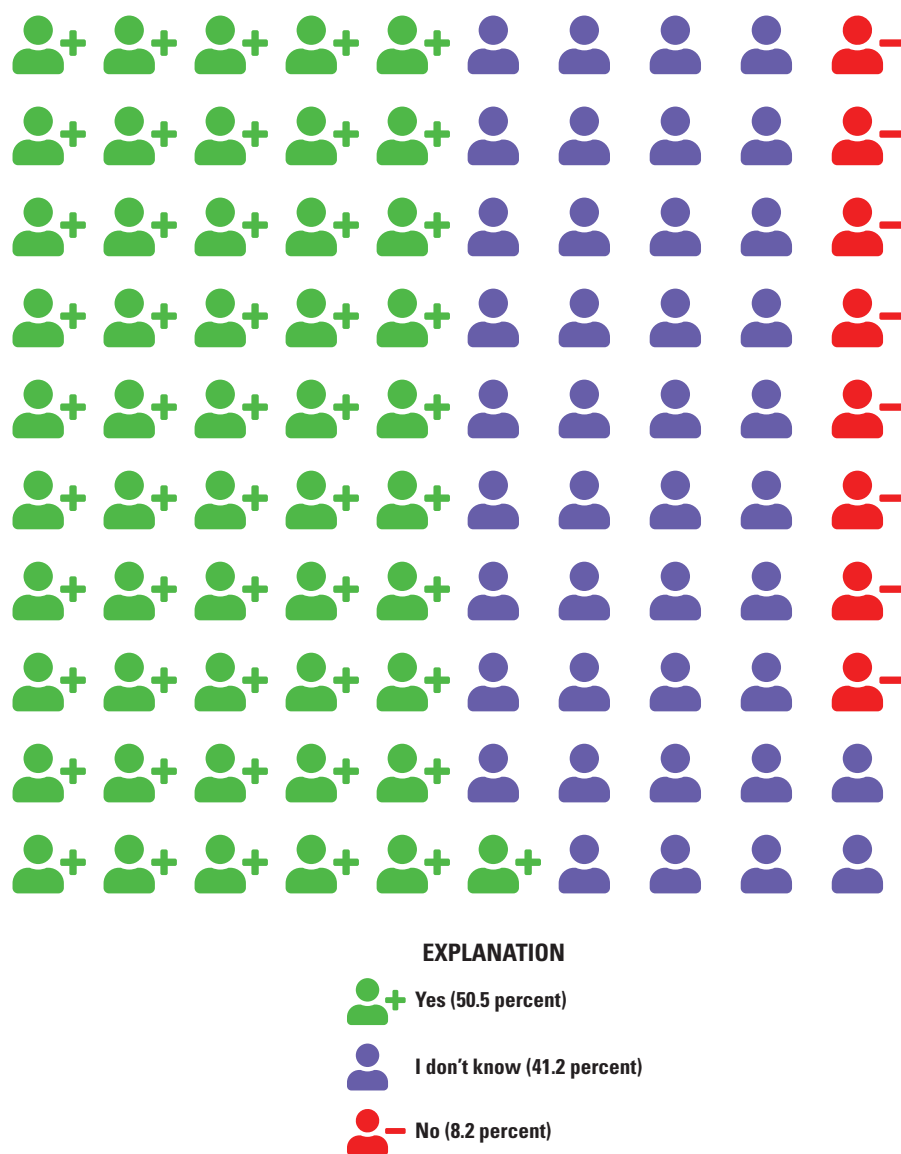
Figure 7. Percentage of researchers reporting that their Center has a documented process for creating project data management plans. Number of responses was equal to 485 (Appendix 1, question 10).

EXPLANATION

Yes (48.2 percent)

I don't know (37.9 percent)

No (13.8 percent)
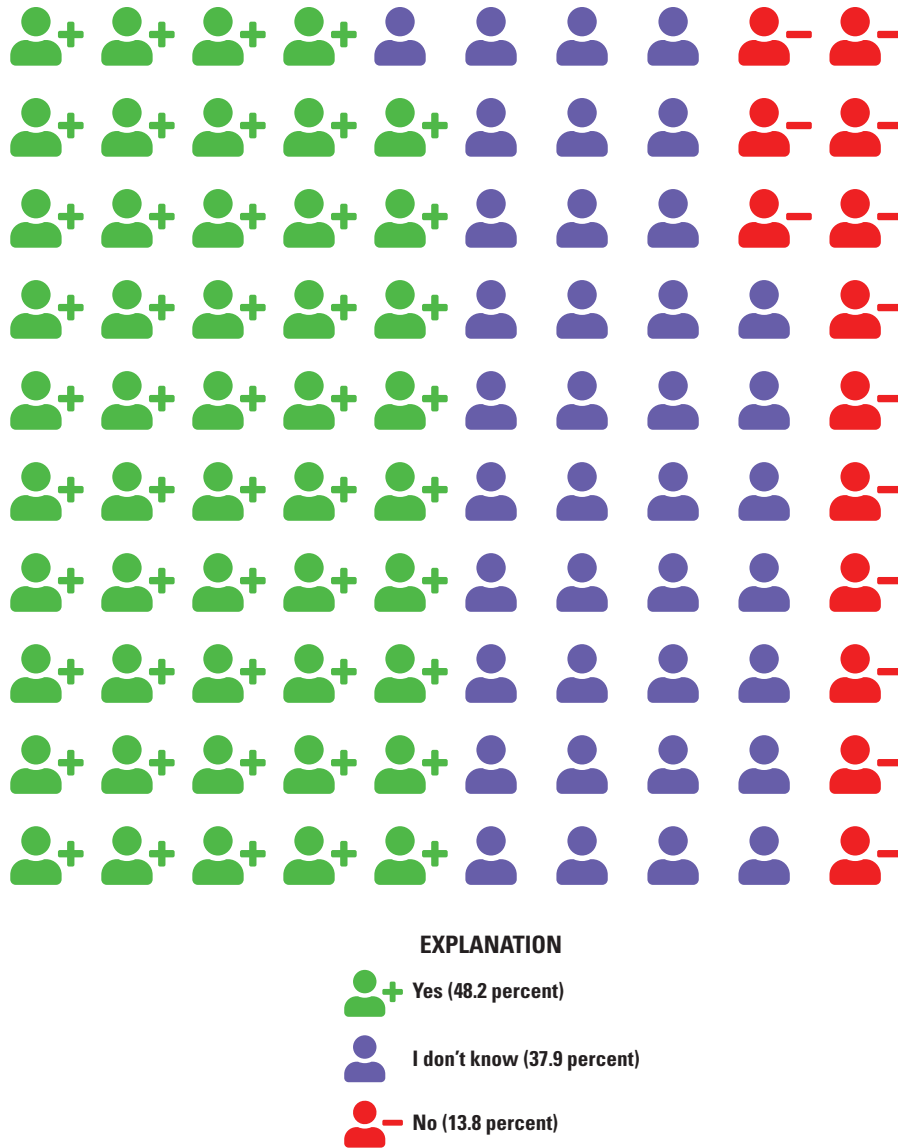
Figure 8.   Percentage of researchers reporting that someone is available at their Center to help them create a data management plan. Number of responses was equal to 485 (Appendix 1, question 12).

Data Management Plan Review and Approval Process

Data managers from 26 Science Centers and Center Directors from 13 Science Centers reported that their Center has a documented process for reviewing and approving DMPs. When these results are combined, they indicate there are a total of 33 unique USGS Science Centers with documented processes for reviewing and approving DMPs. Researchers were not asked about processes for reviewing and approving DMPs because by policy, it does not fall within their area of responsibility.

Directors were not asked about whether their Center has a documented policy for updating DMPs because this knowledge is assumed for those roles.

Even though many researchers were not sure if there is a documented policy for updating DMPs at their Center (fig. 9), approximately one-half of them (49.3 percent) reported project progress reviews do include an update on the status of planned data management activities (fig. 10).

EXPLANATION

I don't know (65.6 percent)

Yes (22.1 percent)

No (12.4 percent)
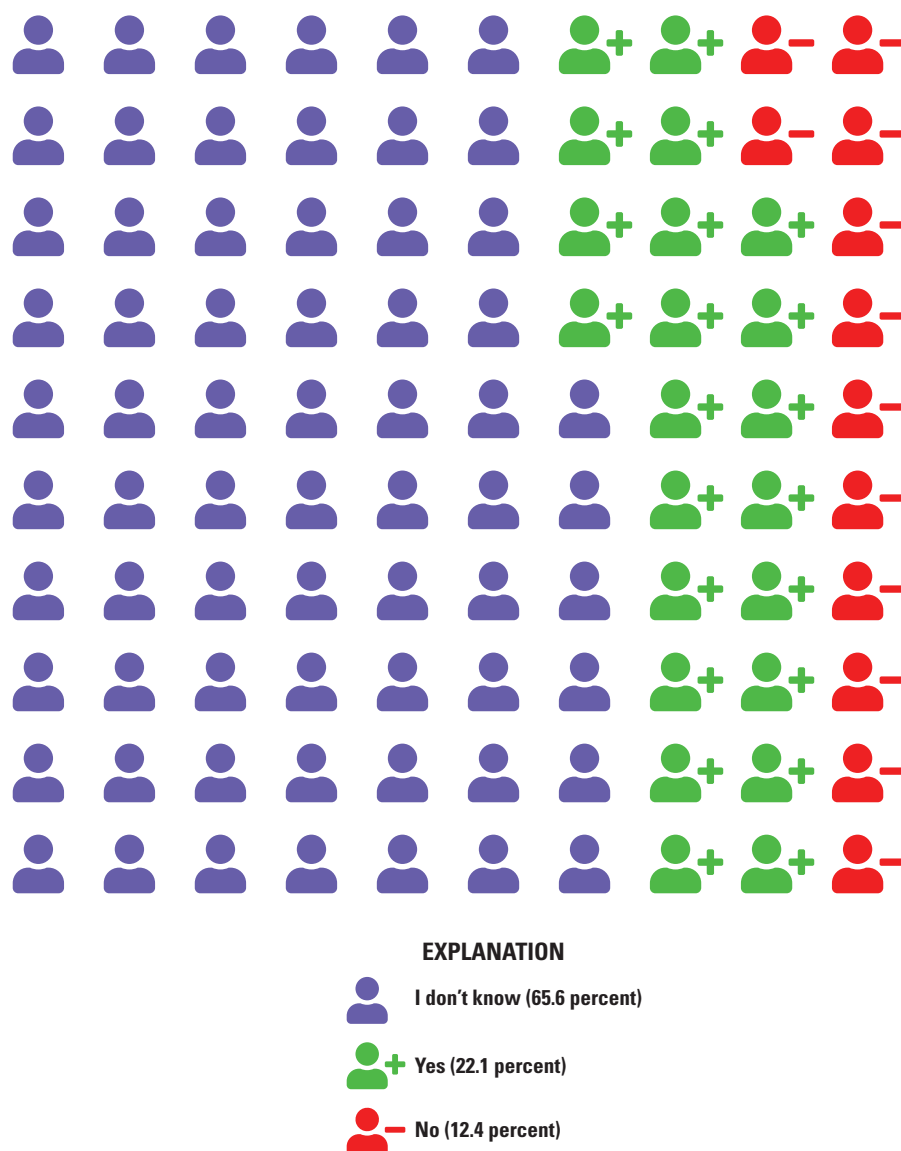
Figure 9.   Percentage of researchers reporting that their Center has a documented policy for updating data management plans. Number of responses was equal to 485 (Appendix 1, question 11).

Variables Related to Data Management Plan Creation

At Centers where data managers reported having a role related to DMPs, researchers were significantly more likely to have reported that they had a DMP for their most recent project, χ^2 (1, number of responses [N] equal to [=] 486) = 9.33, probability value [p] = 0.002 (standard residual = 3.15, p = 0.001). Standard residual is a measure of the strength of the difference between observed and expected values, and p values less than or equal to 0.05 indicate a statistically significant relation between categories. Where Center Directors reported that they are directly involved in review and approval of project DMPs or ensuring they are created, there was no

statistically significant relation with researchers reporting that a DMP was created for their most recent project, χ^2 (1, N = 210) = 0.41, p = 0.520.

There is a statistically significant association between researcher awareness of DMP creation assistance and whether DMPs were created (by the respondent, by a project member, or DMP already in existence), χ^2 (1, N = 485) = 46.47, p = 0.000. There also is a positive association between researchers knowing someone is available to help with DMP creation and DMPs being created (standard residual = 6.91, p = 0.000).

There also is a statistically significant association between researcher awareness of a documented DMP creation process and whether DMPs were created for their most recent project, χ^2 (1, N = 485) = 47.26, p = 0.000, and there is a

EXPLANATION

Yes (49.3 percent)

No (38.3 percent)

I don't know (12.5 percent)
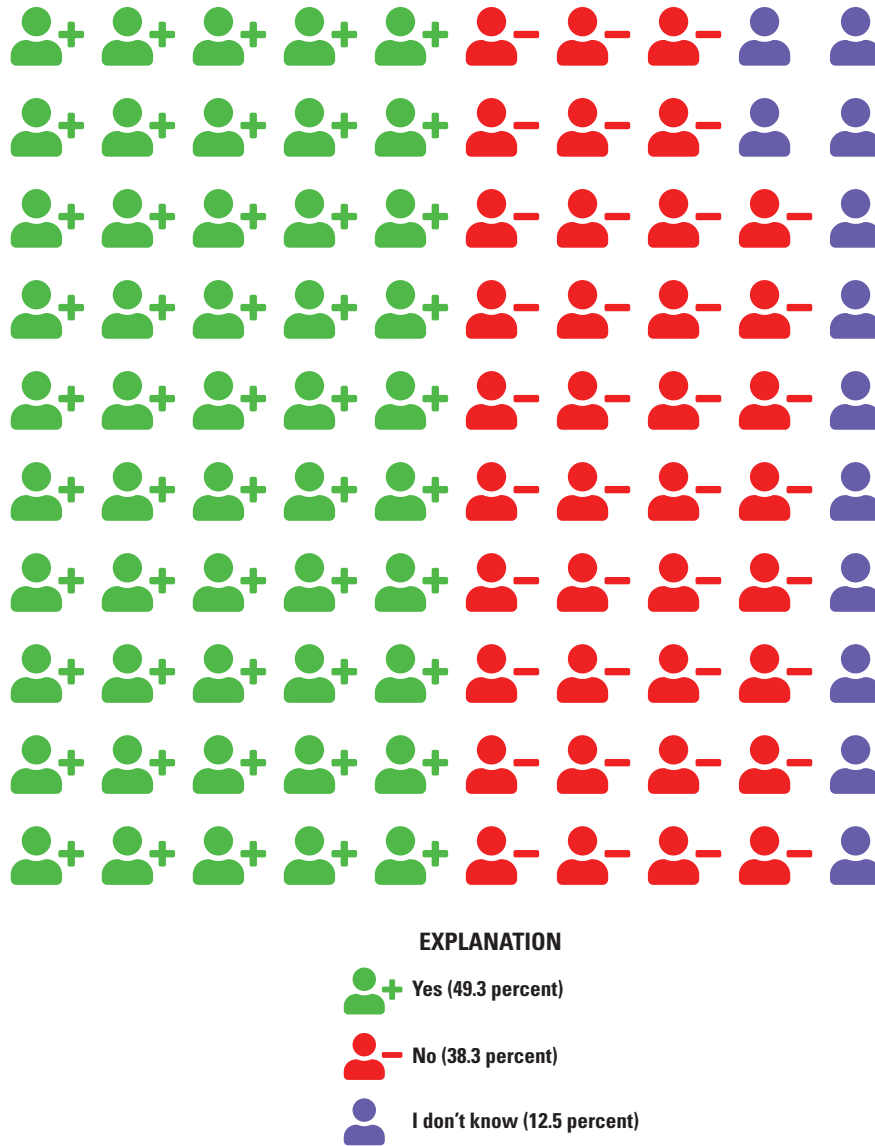
Figure 10. Percentage of researchers reporting that their project progress reports or reviews include an update on the status of planned data management activities. Number of responses was equal to 481 (Appendix 1, question 6).

positive association between knowing about a Center process for creating DMPs and DMPs being created (standard residual = 6.97, p = 0.000).

Data Management Plan Access

Of those researcher respondents who had a DMP for their most recent project, the majority (62.6 percent) reported their project DMP was in a shared folder, drive, or database internally available for their Center; however, almost one-third of researchers (32.4 percent) said that the DMP was only available by request from a project team member's computer (fig. 11). Most data managers (91.9 percent) reported that DMPs for their Center are available in a shared folder, drive,

or database, with only 13.5 percent reporting that DMPs are only available by request from a project team member's computer (fig. 11). Center Directors were not asked this question because they are not directly involved in DMP access and maintenance.

Ideal Location for Data Management Plans

When asked about the ideal location for USGS project DMPs to be stored and accessed, the most common response from researchers (45.0 percent), data managers (65.5 percent), and Center Directors (58.8 percent) was that DMPs should be internally available in a shared folder, drive, or database for their Center (fig. 12).
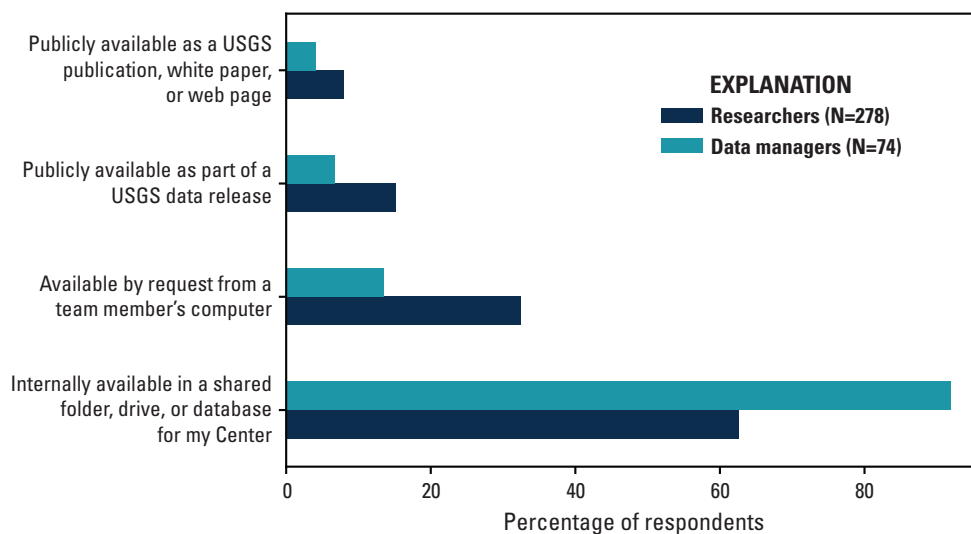
Figure 11. Accessibility of data management plans as reported by researchers (Appendix 1, question 8) and data managers (Appendix 2, question 9). N=, number of responses equal to; USGS, U.S. Geological Survey.
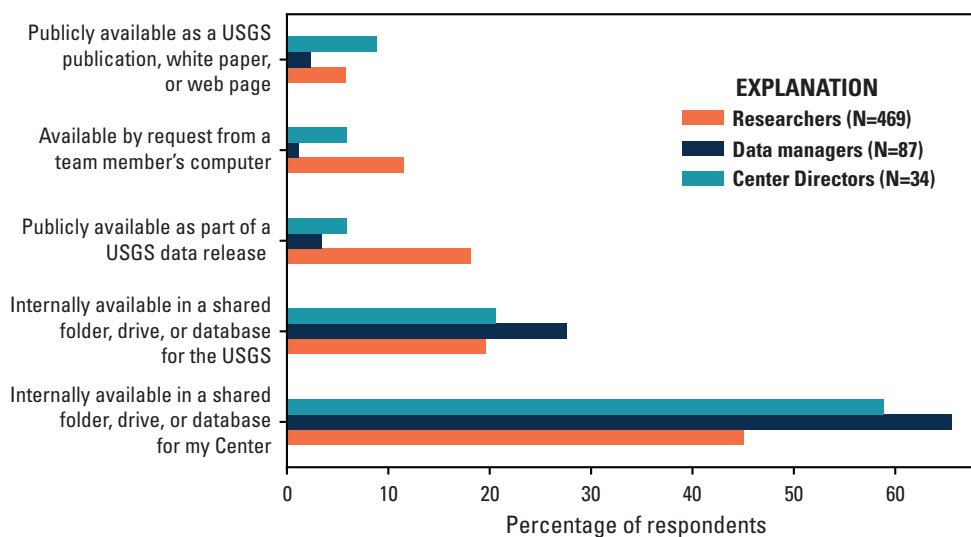
Figure 12. Ideal location for data management plan storage and access as reported by researchers (Appendix 1, Q23), data managers (Appendix 2, Q22), and Center Directors (Appendix 3, question 10). N=, number of responses equal to; USGS, U.S. Geological Survey.

Data Management Plan Maintenance

Researchers were asked how their DMPs are maintained, either as a living document that receives updates throughout the lifecycle of the project or as a static document that is not changed or updated (appendix 1, Q9). More researchers (42.0 percent) reported their DMPs are maintained as living documents (fig. 13).

Similar results were present when data managers were asked the same question regarding how DMPs for their Center are primarily maintained (appendix 2, Q10). More than

one-half (56.0 percent) of data manager respondents indicated their Center's DMPs are maintained as living documents. Less than one-third (30.7 percent) reported DMPs are maintained as static documents, and 13.3 percent reported they did not know how DMPs are maintained. Center Directors were not asked this question because they are not directly involved in maintaining DMPs.

A Pearson's χ^2 test of independence was performed to examine the relation between providing updates about data management during project progress reviews (answer options were "Yes," "No," or "I don't know") and how DMPs are

EXPLANATION

Living document (42.0 percent)

Static document (38.9 percent)

I don't know (19.1 percent)

Figure 13. Researcher responses to whether their data management plans are treated as living documents that are updated throughout projects or as static documents that are not updated (number of responses was equal to 293; Appendix 1, question 9).

maintained (answer options were "Living Document," "Static Document," or "I don't know"). The association between these variables was statistically significant, χ^2 (4, N = 290) = 41.02, p = 0.000. There is a positive association between providing updates about data management during project reviews and maintaining DMPs as living documents (standard residual = 3.93, p = 0.000). There also is a positive association between knowing a policy for updating DMPs exists and maintaining DMPs as living documents, χ^2 (2, N = 293) = 23.69, p = 0.000 (standard residual = 4.79, p = .000).

Researchers were interviewed to explore the idea of maintaining DMPs as living documents and knowledge of DMP update policies. Interviews of researchers who had reported maintaining DMPs as living documents and having

documented DMP update policies in their survey responses revealed that they were often unsure or unaware of policies around updating DMPs, or they disclosed that no formal policy exists. Several interviewees from this group also mentioned they did not update their DMPs but thought it was a good idea or had plans to do so in the future.

Data Management Plan Perceptions

In the second section of the survey, researchers were asked to consider their prior projects when responding to questions related to their perceptions of interactions with DMPs. When asked if they had ever created, or helped to create, a

project DMP (appendix 1, Q13), 59.1 percent of researchers answered "Yes" and 40.9 percent answered "No." Researchers who responded "No" were asked why they had not created a project DMP and were allowed to choose multiple responses from a list, as well as provide free-text responses (appendix 1, Q22). Most of these respondents (62.2 percent) indicated they had never been required to create a DMP for any of their projects, and nearly one-fourth of respondents (23.5 percent) reported someone else had always created the DMP for them. Many free-text responses indicated researchers were not sure about DMP requirements or the purpose of a DMP. For example,

> "I don't know what you mean by a data management plan."
>
> USGS researcher

> "I don't know what a DMP is for."
>
> USGS researcher

Some researchers also noted they publish data releases and articles describing methods and data management activities, indicating the belief that this fulfills the requirements or purpose of a DMP. For example,

> "DMP has never been explained. I do data releases constantly for everything but that is it."
>
> USGS researcher

> "My group produces data releases frequently because our focus is peer reviewed interpretive papers, and DRs [data releases] are required. I am curious if this constitutes an adequate DMP, but I don't think I have done any formally."
>
> USGS researcher

Data Management Plan Motivations and Value

Researchers who had been involved with the creation of DMPs were asked what their main motivations were for creating them (appendix 1, Q14). They were allowed to choose multiple responses from a list (table 3), as well as provide free-text responses. Most of the free-text responses indicated that DMPs were required by the researcher's project funding, both within USGS and external funding sources.

Data managers were also asked what researchers' main motivations are for creating project DMPs (appendix 2, Q14) (table 3). In free-text responses, data managers, like researchers, noted there are requirements other than USGS and Center policies that may motivate DMP creation, such as at the USGS Mission Area level or from external funding sources. One respondent also noted a motivating factor may be the need to document data sharing and ownership details for a project. Three respondents specifically noted IT planning as a motivation. According to the survey results, the least motivating factors for DMP creation were as follows: for researchers,

"DMPs help me learn about new data management practices and requirements" (15 percent), and for data managers, "DMPs are required by my Center Director" (44.3 percent).

Center Directors were asked a slightly different version of the question about DMP motivation: "What do you think is the main value(s) of data management plans (DMPs)?" (appendix 3, Q1). Like most data managers, most Center Directors indicated that the primary value of DMPs is to help researchers define how they will manage and track their data throughout their projects, followed by helping Centers plan for resources needed for managing project data (table 4). For Center Directors, the least selected option in response to the question was "DMPs help project teams communicate with project partners" (29.4 percent).

Six free-text responses were also submitted by Center Directors in response to this question. One-half of these responses related to ensuring data are appropriately handled for long-term storage and use. One respondent reported DMPs help researchers communicate with their Center's database management team early (for example, at the beginning of a project), and another respondent noted that DMPs help identify potential data management issues before starting the project.

Interview responses emphasized many researchers and data managers feel that buy-in and compliance with DMP policies and the creation of DMPs would be increased if Center Directors or other executive leaders emphasized the importance of DMPs. For example,

> "But I think one of the biggest problems we've come across is not having, you know, buy-in completely from Center Management and a level up, you know, it's like they [DMPs] were a great idea and everyone was behind them and then it's like OK. These really aren't enforceable * * *"
>
> USGS data manager

Data Management Plan Challenges

In each survey, participants (researchers, data managers, and Center Directors) were asked to select all the challenges that they face when creating project DMPs, as well as to indicate their greatest challenge (appendix 1, Q19; appendix 2, Q19; appendix 3, Q6). The two challenges that were selected by most researchers and data managers were that no one ever reads the DMP and DMPs take too much time to create (fig. 14, fig. 15). The two challenges selected by the most Center Directors were that a DMP tracking mechanism is needed and that they don't have a formal process for data management planning (fig. 16).

The top challenge selected by most researchers, data managers, and Center Directors was that DMPs take too much time. The second most common response for researchers was "I don't face any challenges," followed by "No one ever reads it." (fig. 14). The second most common response for data managers was "we don't have a formal process for

Table 3.   Researcher (286 total) and data manager (88 total) responses to the multiple-selection question regarding motivations for creating project data management plans (DMPs).

[%, percent]

Multiple-selection options	Researchers (%)	Data managers (%)
DMPs are required by U.S. Geological Survey policy	65.0	72.7
DMPs are required by my/our Center Director	50.3	44.3
DMPs help me/researchers define how I/they will manage my data throughout the lifecycle of my/their projects	44.8	78.4
DMPs help me/Centers plan for resources (staff time and project funding) needed for managing project data	26.6	67.0
DMPs help me/researchers improve the quality of my/their research and data products	25.2	53.4
DMPs help me/project teams communicate information with project evaluators and data managers at my Center	24.8	61.4
DMPs help me/project teams communicate with project partners	23.4	45.5
DMPs help me/researchers learn about new data management practices and requirements	15.0	45.5

Table 4.   Center Director responses (34 total) to the multiple-selection question, "What do you think is the main value(s) of data management plans (DMPs)?"

[%, percent]

Multiple-selection options	Center Directors (%)
DMPs help researchers define how they will manage and track their data throughout their projects	91.2
DMPs help Centers plan for resources (staff time and project funding) needed for managing project data	73.5
DMPs help improve the quality of research and data products	55.9
DMPs help project teams communicate information with project evaluators and data managers	44.1
DMPs help researchers learn about new data management practices and requirements	44.1
DMPs help project teams communicate with project partners	29.4

data management planning," followed by "No one ever reads it." (fig. 15). The second most common response for Center Directors was a three-way tie among the need for a DMP tracking mechanism, tools and templates, and a formal process for data management planning (fig. 16).

Researchers also submitted many other challenges to DMP creation as free-text responses. Some of these responses indicated that researchers need more assistance when it comes to creating DMPs. For example,

> "I could use an experienced facilitator or clear templates to help create the DMP."
>
>   USGS researcher

Some researchers also noted that current tools and templates (as of March 2021) are "clunky," and they feel uncertain how to respond to some of the questions in DMP templates. Other respondents described there is uncertainty at the beginning of the project, which makes them wary of putting something in writing that they may not be able to deliver.

Data managers also included free-text responses when asked about the various challenges faced when creating DMPs. Several respondents reported the value of DMPs is unclear and they are an "administrative burden on researchers." Data

manager respondents also indicated there is "no Center leadership support" for DMPs and "no accountability from Center management for completing or updating a DMP." Other respondents noted the DMP templates they have used in the past do "not always fit the project."

When asked about the top challenges faced when creating or encouraging researchers to create DMPs, data managers noted in their free-text responses that USGS researchers "think that [DMPs] have little benefit" and may "need a few cycles of reviews * * * to embrace [the] benefit of DMPs." Other respondents also noted there is "little guidance on how to complete a DMP, where the DMP templates can be found, or how to view the inventory of already established DMPs."

Center Directors reported other challenges, including a poor understanding of data management and DMPs among project leads contributing to their reluctance to create them. Additionally, they reported a lack of understanding among project leads of records management tasks, such as archiving, to properly complete and execute DMPs. They also noted that, when created, DMPs are not updated or followed.
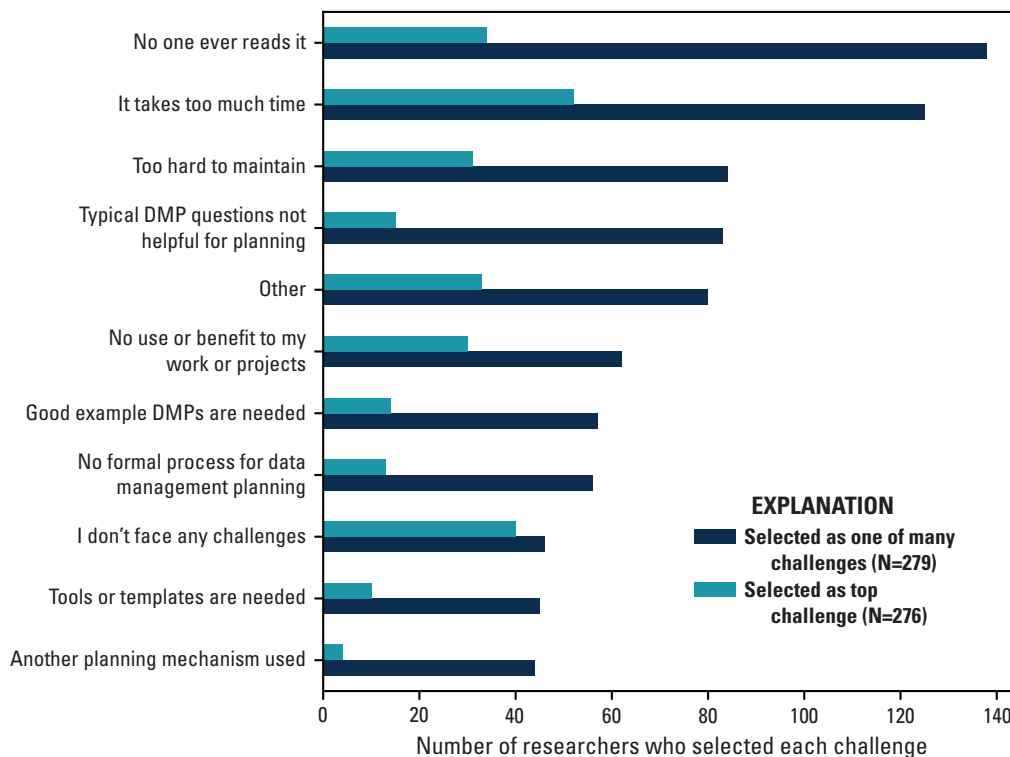
Figure 14. Challenges researchers face related to data management plan (DMP) creation and use (Appendix 1, questions 18 and 19). N=, number of responses equal to.
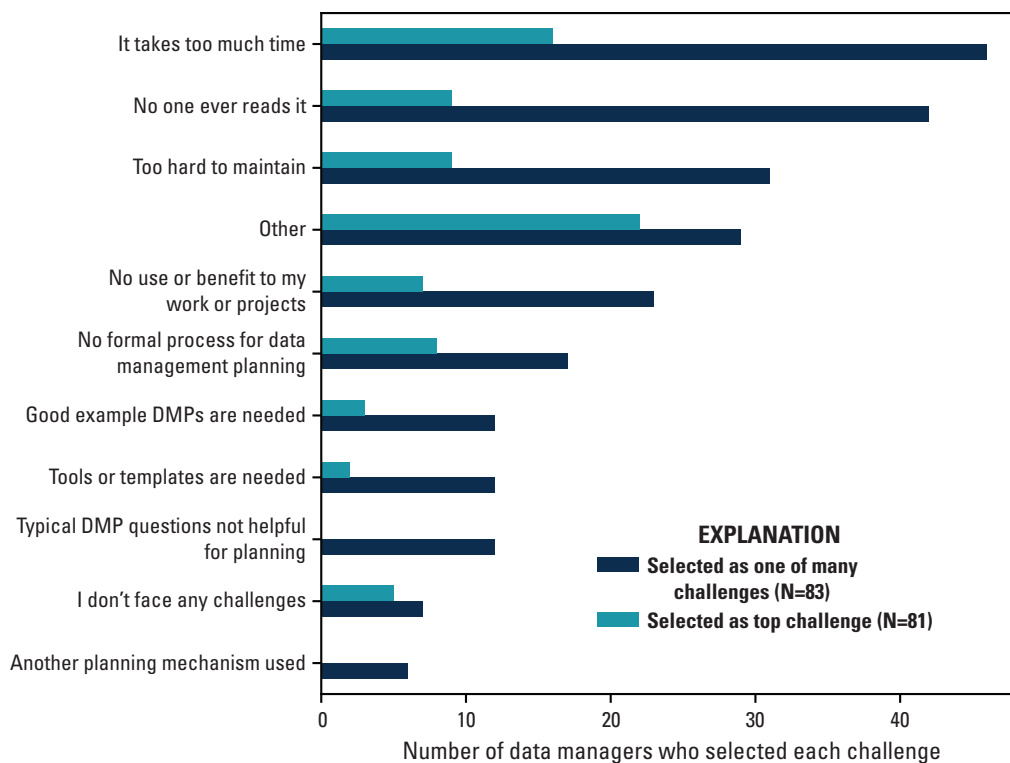
Figure 15. Challenges data managers face when creating or encouraging a researcher to create a project data management plan (DMP; Appendix 2, questions 18 and 19). N=, number of responses equal to.
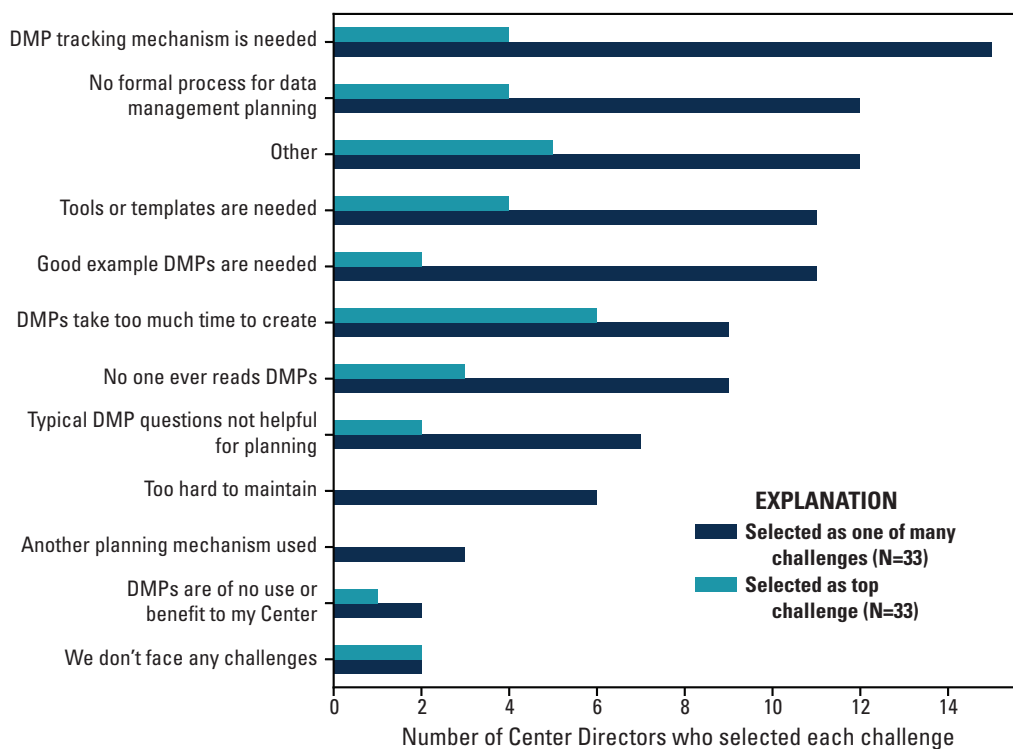
Figure 16. Challenges Science Centers face with respect to creating and using data management plans (DMPs) as reported by Science Center Directors (Appendix 3, questions 5 and 6). N=, number of responses equal to.

Data Management Plan Usefulness

The following sections describe the survey results that are related to DMP usefulness. These results include who DMPs are most useful for, how they are useful, what elements are most useful, what variables might be influencing the perceived usefulness of DMPs, and respondents' suggestions for making them more useful.

Usefulness by Role

Researchers, data managers, and Center Directors were asked if DMPs were useful to their work (appendix 1, Q15; appendix 2, Q15, appendix 3, Q2), and how useful they thought DMPs are to people in other roles (appendix 1, Q20; appendix 2, Q20, appendix 3, Q8). Available responses to these questions were presented as Likert scales (Joshi and others, 2015). In general, all respondents believe DMPs are most useful for data managers (figs. 17, 18, and 19). Most researchers (fig. 17) reported DMPs were less useful to themselves and others, compared to data managers (fig. 18) and Center Directors (fig. 19), most of whom reported DMPs were somewhat or very useful to themselves and others.

How Data Management Plans are Useful

Researchers, data managers, and Center Directors were asked to describe the ways that DMPs are useful to their projects or work (appendix 1, Q16; appendix 2, Q16; and appendix 3, Q3). Many of the researcher's free-text responses revolved around planning for data release and archiving data at the end of projects, as well as ensuring there is awareness of FSP requirements. For example,

> "We have had to figure out where our data will be published if they don't fit into NWIS [National Water Information System]—Science Base is where our team is putting those data."
>
> USGS researcher

Data manager and Center Director responses echoed this theme, noting DMPs help to catch potential oversights that may cause issues later in a project, especially related to FSP. For example,

> "Many * * * project chiefs don't know all there is to know about USGS data collection requirements * * *—leading to FSP issues. The DMP reviews during proposal review process catches these issues early."
>
> USGS data manager

> "[DMPs] also help connect researchers with our database management and support team early in
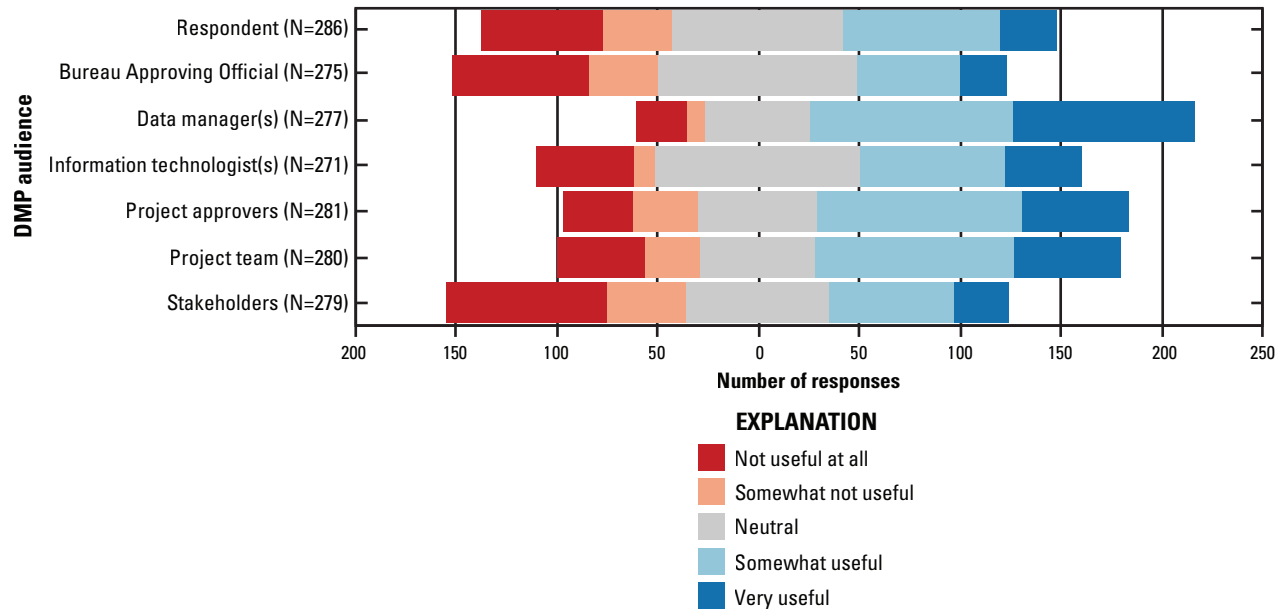
Figure 17.   Responses to Likert scale (Joshi and others, 2015) questions about the usefulness of data management plans (DMPs) to researcher respondents' own work and how useful they think data management plans are to people in other roles. Each question was optional, so the number of responses for each question varies (Appendix 1, questions 15 and 20). N=, number of responses equal to.

Figure 18.   Responses to Likert scale questions about the usefulness of data management plans to data manager respondents' own work and how useful they think data management plans (DMPs) are to people in other roles. Each Likert scale question was optional, so the number of responses for each question varies (Appendix 2, questions 15 and 20). N=, number of responses equal to.
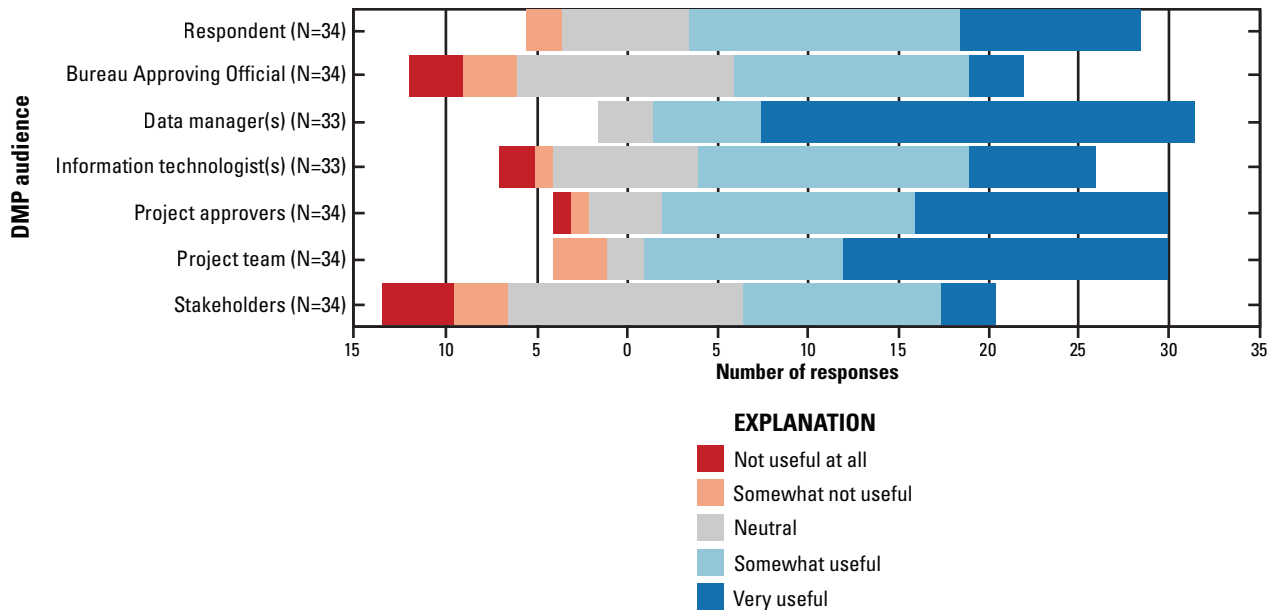
Figure 19.   Responses to Likert scale questions about the usefulness of data management plans to Center Director respondents' own work and how useful they think data management plans (DMPs) are to people in other roles. Each Likert scale question was optional, so the number of responses for each question varies (Appendix 3, questions 2 and 8). N=, number of responses equal to.

the project life cycle which helps prevent problems later on."

    USGS Center Director

Researchers, data managers, and Center Directors also noted DMPs help ensure they have budgeted time, money, and staff to complete data release and archiving activities. For example,

"[A DMP] Allows me to set aside time/resources to ensure that data management is occurring."

    USGS researcher

"[DMPs are a] very important planning tool. DMPs force management to understand how much time will actually be required to complete the project."

    USGS data manager

"[A DMP] Helps project chiefs think through where all their data are going at the inception of a project, so they can properly budget or estimate staff time for generating data releases and metadata."

    USGS Center Director

In many cases, respondents reported DMPs have been helpful for outlining data roles, responsibilities, and agreements, including documenting when someone is not responsible for releasing data.

"[A DMP] Serves as a disclaimer for data that I shouldn't be responsible for."

    USGS researcher

"DMPs help to get [sic] ensure that all personnel responsible for the data are identified and kept informed from initial creation through project completion and that the necessary resources are planned for."

    USGS data manager

"DMPs can make investigators more aware of the need to plan for the final disposition of data, which can involve more advanced planning about ownership of data, data sharing agreements, nondisclosure agreements, etc."

    USGS Center Director

Although many of the responses revolved around data management processes at the end of the data lifecycle, such as publishing, there were examples of DMPs being useful for processes closer to the beginning of the data lifecycle, such as acquiring, storing, and organizing data.

"Developing the DMP requires thinking about what data will be collected, how it will be organized, and where it will be stored in advance of the onset of data collection."

    USGS researcher

Understanding the anticipated storage size for IT infrastructure purposes was of particular interest to data managers and one Center Director.

"[DMPs have been useful] in planning with our IT department for storage of particularly large datasets, such as imagery."

USGS data manager

"[DMPs provide] some understanding of local server needs and requirements."

USGS Center Director

In the free-text question about DMP usefulness, 47 of the 214 researcher respondents (22.0 percent) described why they felt DMPs were a waste of time and not useful to them.

"My research group already has strict data management procedures, processes, and detailed SOPs, so the additional USGS required DMPs do not have much added benefit for my group."

USGS researcher

Within these responses, there were many statements about how DMPs are only useful for people besides the respondent, such as new researchers, researchers with large projects or datasets, or people in different roles.

"* * * [DMPs are] useful for new * * * scientists. * * * Those of us who have been doing science and research for many years know how to manage data and it is an intrinsic part of what we do."

USGS researcher

"Usually, I'm the only one doing the work and have a good sense of where data are or where products will be published."

USGS researcher

"I feel it's probably more useful for IT to receive a heads up for projects for which large amounts of disk or server space may be required."

USGS researcher

Data Management Plan Element Usefulness

Researchers, data managers, and Center Directors were asked how useful various elements or content sections within a DMP are to their work (figs. 20–22). Researchers, data managers, and Center Directors rated the following as the top three most useful elements that may be included in a DMP:

1. Where the data will be stored and backed up;

2. Plans for data preservation and disposition; and

3. Where the data will be released.
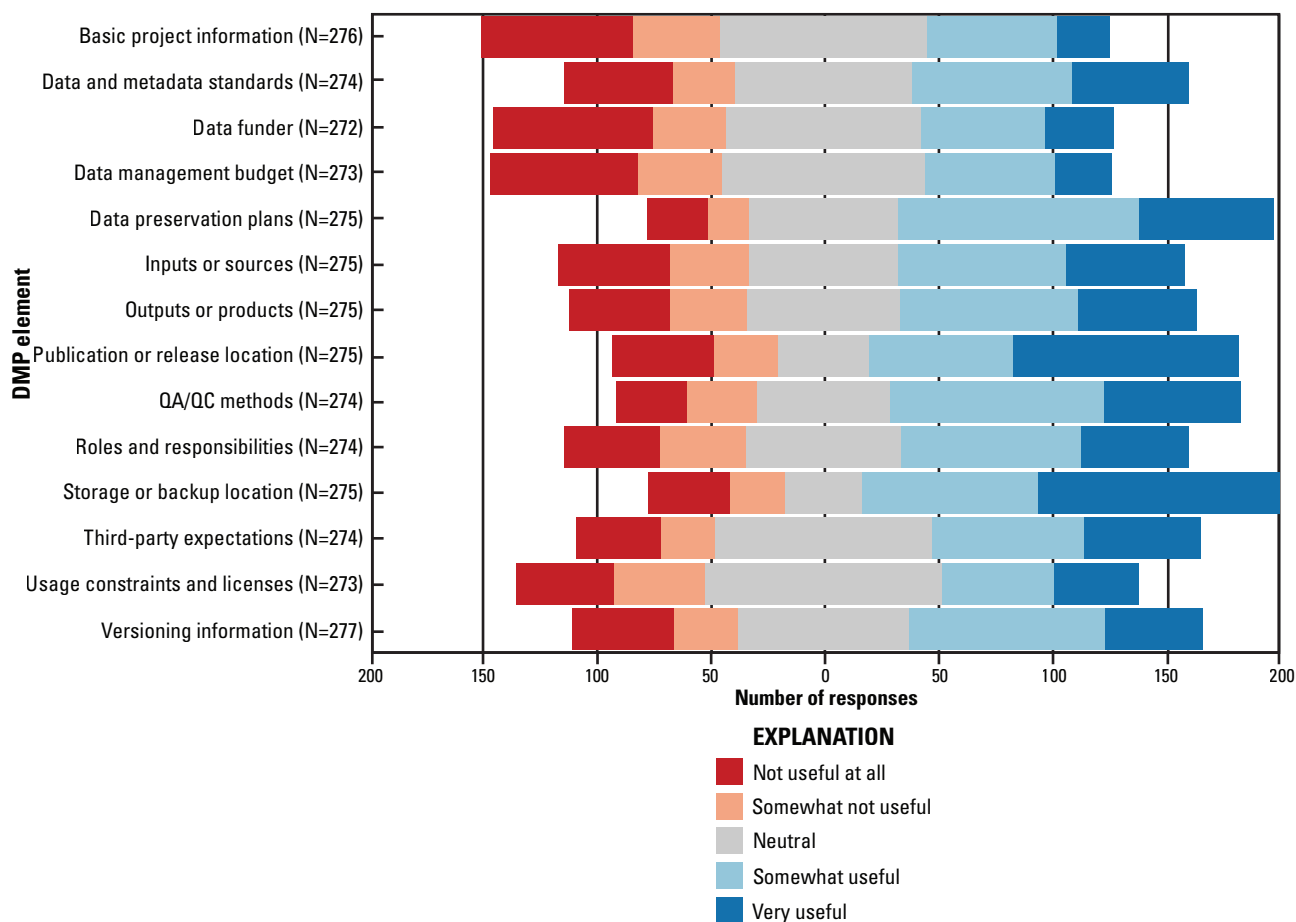
Figure 20. Researcher responses to Likert scale questions about the usefulness of data management plan (DMP) elements. Each Likert scale question was optional, so the number of responses for each question varies (Appendix 1, question 21). N=, number of responses equal to; QA/QC, quality assurance and quality control.
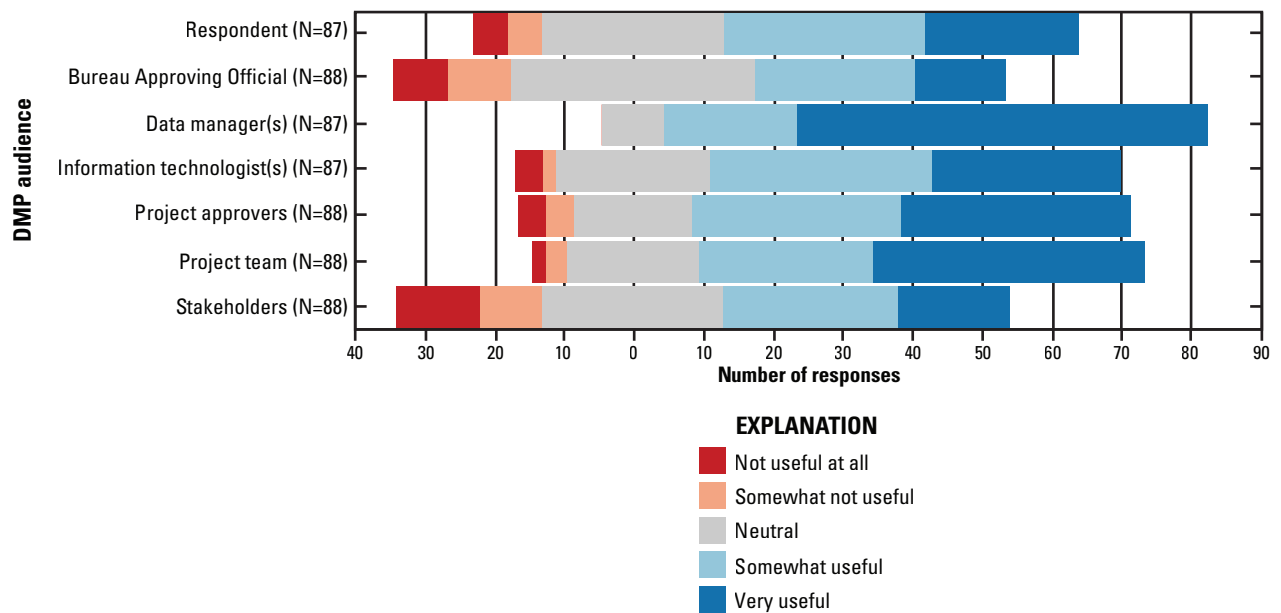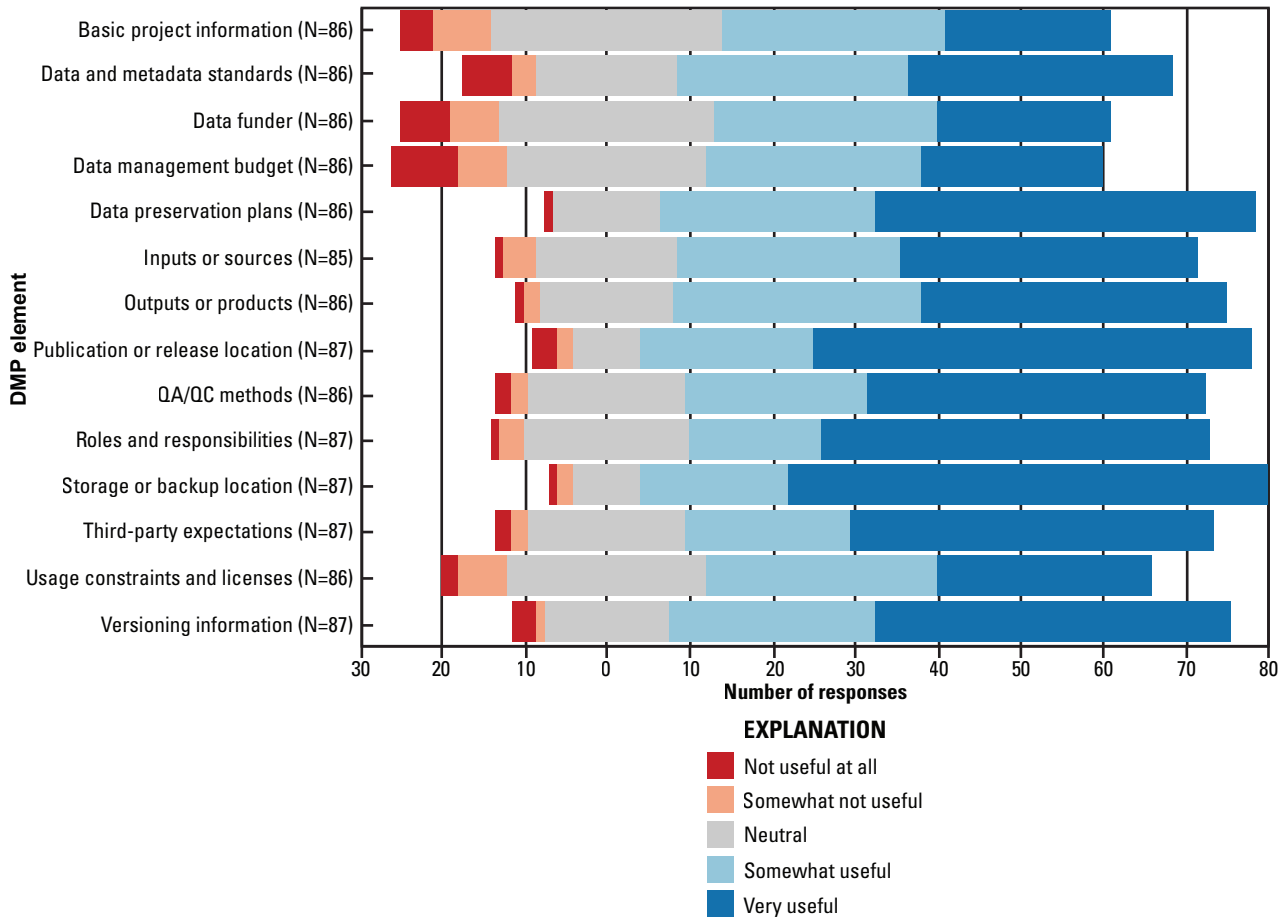
Figure 21. Data manager responses to Likert scale questions about the usefulness of data management plan (DMP) elements. Each Likert scale question was optional, so the number of responses for each question varies (Appendix 2, question 21). N=, number of responses equal to; QA/QC, quality assurance and quality control.

Data managers indicated most of the elements in a DMP are relatively equally important, whereas researcher and Center Director responses showed more variation in the level of importance across elements.

Variables Related to Data Management Plan Usefulness

A series of Pearson's χ^2 tests of independence were performed to test which, if any, responses to other questions were associated with researcher respondents reporting that DMPs were useful to their work. The variables that were tested against DMP usefulness were as follows:

- Whether their Center has a documented process for creating DMPs;

- Whether their Center has a documented policy for updating DMPs;

- Whether someone is available at their Center to help with DMP creation;

- How DMPs are maintained; and

- Whether project progress reviews include updates on planned data management activities.

There is no statistically significant association (p greater than 0.05) between having a process or policy for creating or updating DMPs and how useful researchers find DMPs to be to their work, χ^2 (2, N = 286) = 3.77, p = 0.152 and χ^2 (2, N = 286) = 5.28, p = 0.071, respectively. However, there are statistically significant associations (p less than or equal to 0.05) between whether someone is available to assist with DMP creation, how DMPs are maintained, and whether project progress reviews include updates on planned data management activities and how useful researchers find DMPs to be to their work.

According to the survey responses, whether someone is available to help with DMP creation is significantly associated with DMP usefulness, χ^2 (2, N = 286) = 7.24, p = 0.027). There is a statistically significant negative association between someone being available to help with DMP creation and researchers reporting that DMPs are not useful at all or somewhat not useful (standard residual = –2.49, p = 0.006).
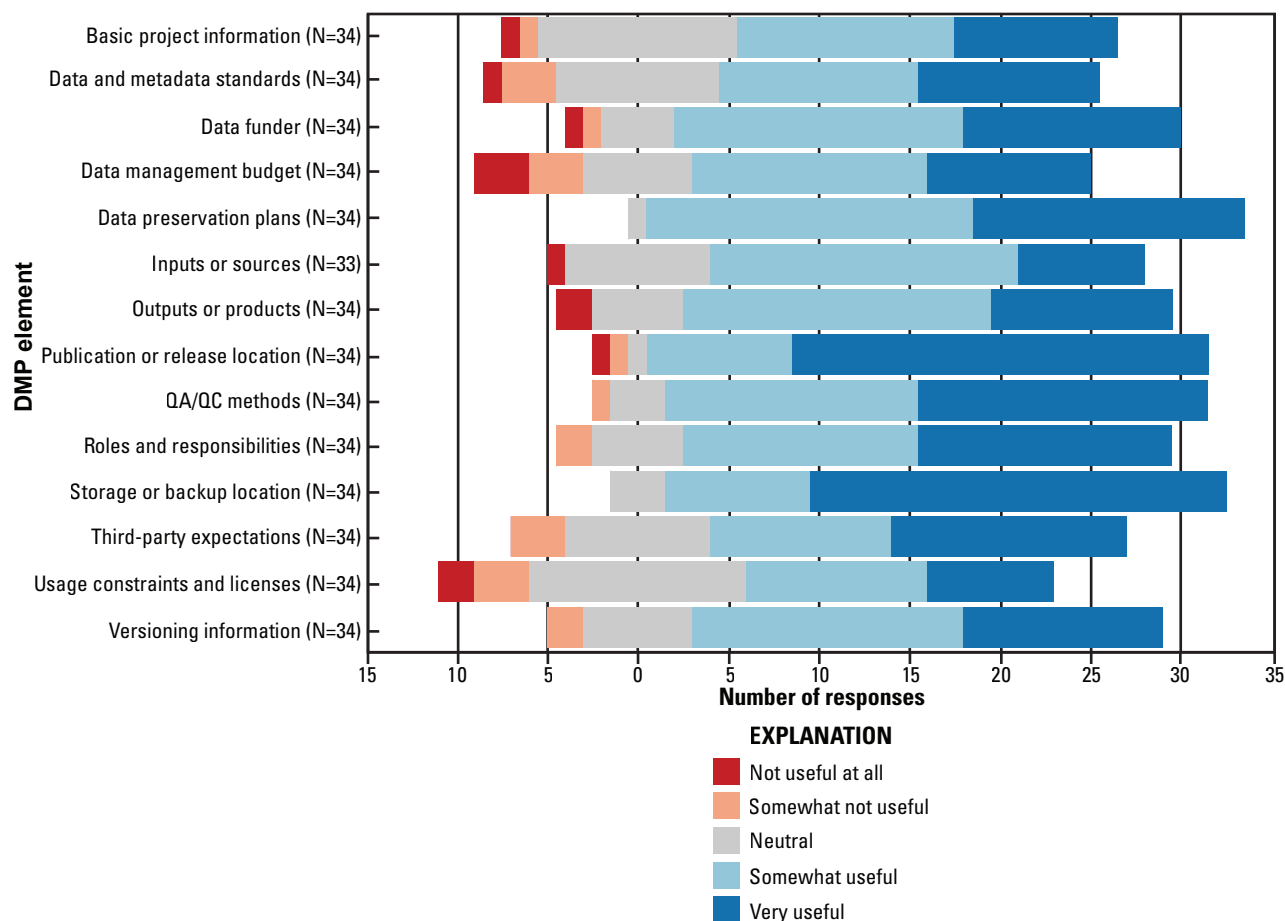
Figure 22. Center Director responses to Likert scale questions about the usefulness of data management plan (DMP) elements (Appendix 3, question 9). N=, number of responses equal to; QA/QC, quality assurance and quality control.

The way DMPs are maintained also is significantly associated with DMP usefulness, χ^2 (4, N = 221) = 27.15, p = 0.000. Maintaining DMPs as living documents (those that are updated) is significantly associated with researchers reporting DMPs are very or somewhat useful (standard residual = 4.98, p = 0.000). In addition, the inclusion of data management updates during project progress reviews also is significantly associated with DMP usefulness, χ^2 (4, N = 283) = 12.79, p = 0.012). Providing updates about data management activities during project progress reviews is significantly associated with researchers reporting the DMPs are very or somewhat useful (standard residual = 3.20, p = 0.001).

Data managers indicated that DMPs are very or some-what useful in more than 82 percent of responses; therefore, Pearson's χ^2 tests of independence were not statistically significant in determining factors related to DMP usefulness for this survey group. Given the comparatively small sample size of Center Director responses (34), nearly three-fourths of which indicated that DMPs are very or somewhat useful, Pearson's χ^2 tests of independence were also not statistically significant in determining factors related to DMP usefulness for this survey group.

Project Delays

Researchers, data managers, and Center Directors were asked if they have encountered challenges or delays at the end of a project and if earlier planning and documentation could have improved or mitigated those delays (appendix 1, Q28; appendix 2, Q28, appendix 3, Q22).

Causes of delays commonly described in free-text responses to this question were related to unawareness that a data release was required by USGS policy and uncertainty on how to complete a data release or how much time it would take to release data.

"After manuscripts are approved by the Bureau Approval Official, I am informed a Data Release is required before moving forward. I wish I was told that earlier in the process to make concurrent dis-semination. I wouldn't mind doing it, but it was a surprise when I thought everything was done."

USGS researcher

Data managers noted researchers they worked with were either "surprised" they had to make their data public or that starting the data release process was often left to the last minute. They also noted that the benefits of better documentation throughout the lifecycle of a project was only understood after researchers began the data release process. One Center Director noted failure to plan for the earlier stages in the data lifecycle, such as acquisition, has caused project delays:

> "We've had a few occasions over the years where someone gets out in the field with a new acquisition system and they have not thought through how the data will be curated, processed, and disseminated, or who will be able to do the work and how much time it will take."
>
> USGS Center Director

Additionally, researchers cited a need for help with completing data releases.

> "We do not have adequate support to help organize and release the data we collect."
>
> USGS researcher

Other common causes of delay mentioned by researchers are time-consuming reviews that hold up the release of data and publications and that they didn't properly budget for data management activities. For example,

> "* * * the data go through so many reviews, lots of nit-picking, etc., to the point that the data are 1) not quickly released, and 2) hinder significant publications."
>
> USGS researcher

> "* * * [Project delays were caused by the] lack of budget for data cleanup and metadata [generation], lack of plan for timely release of metadata."
>
> USGS researcher

Researchers and Center Directors reported some challenges were the result of evolving policies or projects. For example,

> "The ever increasing burden of changing rules makes it impossible to prepare ahead for these products."
>
> USGS researcher

> "[A project delay] occurs when the project needs evolve after initiation of the project, which happens because we don't always know what will be found during the investigation."
>
> USGS Center Director

When asked about the ways in which DMPs are useful, (appendix 1, Q16), some researcher respondents described DMPs as being useful for establishing who will be responsible for releasing data at the end of the project. This specific function of DMPs was also mentioned in response to the question regarding project delays (appendix 1, Q28; appendix 2, Q28; appendix 3, Q22) four times by researchers, seven times by data managers, and two times by Center Directors.

> "The lack of a data sharing agreement or DMP outlining data responsibilities really hindered their ability to get responses from the data producer."
>
> USGS data manager

> "It is extremely common that data ownership issues are not properly resolved at the beginning of a project, and this becomes an issue when it's time to release data. Also, issues of data that needs [sic] to be protected or not disclosed is [sic] often not identified at the beginning of the project."
>
> USGS Center Director

Suggestions for Making Data Management Plans More Useful

The goal of the next set of survey questions was to determine how to make DMPs more useful. Researchers were asked how DMPs could be more useful to their projects (appendix 1, Q17). Data managers (appendix 2, Q17) and Center Directors (appendix 3, Q4) were asked how DMPs could be more useful to their work.

The two main suggestions from researchers for making DMPs more useful were for (1) more support from experts to create DMPs and (2) better training. For example,

> "Having a DMP coordinator or some other designated person who is familiar enough with both field and laboratory projects that they could recommend the best forms of data management that meet USGS requirements (and have templates) so that we could efficiently document from the start of any project that we were following USGS policy and best science practices."
>
> USGS researcher

Several researchers and a data manager also noted that DMP tools and templates could be more educational and indicate common data management practices that researchers could choose to implement. For example,

> "The DMPs [sic] templates I have used in the past (as I remember them) ask if you plan to collect data and how you think you will archive that data. My answers were 'yes', and 'NWIS [National Water Information System] or ScienceBase'. If DMPs listed all the types of data our Center usually collects, and lists where that type of data is [sic] usually archived, and then provides links detailing the steps 'how-to' archive that data -- that would be useful."
>
> USGS researcher

"We could use them to more effectively educate scientists on the data archival, release and metadata work flows so that those steps are better planned for and initiated, making them more useful for subsequent users."

> USGS data manager

Another common theme in researcher responses was allowing DMPs to be more simplistic and customizable. Some researchers mentioned being forced into a specific template creates duplicative work because they may already capture this type of information in other places. For example,

"Provide more flexibility so the required DMP and the 'real' DMP were more aligned."

> USGS researcher

Fewer researchers (14) noted that having more standardization in DMPs would be an improvement, compared to those requesting more flexibility (29) in DMPs. In contrast, data managers more often mentioned needing more standardization (15) than more flexibility (2) in DMPs. For example,

"[DMPs] would be more useful if they were consistently prepared and mainted [sic] according to a documented process."

> USGS data manager

Data manager and Center Director responses contained many suggestions for how DMPs could be more useful to their work. These included use and improvement of DMP templates, improved communication within a Science Center, buy-in from Center staff, and established processes for reviewing and updating DMPs. For example,

"If we had a system that was supported from the top of our Center down where DMP creation was enforced, a single storage location was supported, and sharing of DMP information was enabled then I think DMPs could be very useful for my work."

> USGS data manager

"If we do a better job of reviewing [DMPs] throughout the project's life and making changes when appropriate to reflect the reality of the work (not just what was planned in the early stages), they will be more helpful."

> USGS data manager

"[DMPs should be] Easier to update, in a format that isn't just a static document, created in a way that is useful but still generic, ways to encourage PIs to update their DMPs, actionable DMPs."

> USGS data manager

"I think there is a fundamental lack of understanding of what a DMP does for a project that needs to be communicated."

> USGS data manager

"DMPs are typically completed by PIs out of obligation, not out of perceived necessity or usefulness. DMPs would be more useful to my work if PI's [principal investigators] had a "carrot" or incentive to complete them and keep them updated throughout the lifecycle of their project."

> USGS data manager

"We have evolved to providing a DMP template, which has proven useful as a tool to get proposers to better articulate their data management practices. It also reduces the effort required by reviewers to assess the plans."

> USGS Center Director

"A Centerwide initiative on the use of DMPs, and dedicating the resources to execute them, is probably needed."

> USGS Center Director

"For complex data collection, I'd like the DMP to trigger a conversation about data management with our data team."

> USGS Center Director

Finally, some BAO respondents who took the Program Coordinator and BAO survey had an additional suggestion for improving the usefulness of DMPs that was not captured within the responses from the researcher, data manager, nor Center Director surveys. Bureau Approving Officials noted that DMPs would be more useful if the BAOs had access to them at the time they are reviewing final manuscripts.

Policy and Resource Awareness

Researchers and data managers were asked if DMPs are required by USGS policy (appendix 1, Q24, appendix 2, Q23). Most researchers (61.7 percent) and data managers (84.3 percent) answered that project DMPs are required, but approximately one-third of researchers (36.9 percent) and fewer data managers (15.7 percent) answered that they did not know if DMPs are required. A small portion of researchers (1.4 percent) and no data managers answered that project DMPs are not required.

Researchers, data managers, and Center Directors were asked which resources they are familiar with related to the creation of DMPs (appendix 1, Q25; appendix 2, Q24; appendix 3, Q21; table 5). The USGS Data Management website resource was selected the most by all three of the survey groups.

In followup interviews, several data managers expressed a need for a primary, common location for resources related to data management and DMPs. Although many data managers knew where to find resources for DMPs and other data

Table 5.   Percentage of researchers, data managers, and Center Directors who indicated they were familiar with data management plan (DMP) resources at the time of completing the survey.

[The specific questions asked can be found in Appendix 1, question 25 for researchers; Appendix 2, question 24 for data managers; and Appendix 3, question 21 for Center Directors. A total of 243, 89, and 33 responses were recorded for researchers, data managers, and Center Directors, respectively. NA indicates the option was not available on the survey. %, percent]

Resource options	Researchers (%)	Data managers (%)	Center Directors (%)
U.S. Geological Survey Survey Manual chapter 502.6	NA	NA	78.8
U.S. Geological Survey Data Management website (https://www.usgs.gov/data-management)	40.8	77.5	75.8
Center's shared resources (for example, SharePoint site)	35.2	48.3	69.7
Center or program DMP template	32.1	65.2	66.7
DMPTool.org	6.6	30.3	21.2
DMPEditor	0.2	1.1	0.0
Not familiar with any DMP resources	31.9	7.9	3.0

management tasks, they mentioned that the resources were widespread across several platforms, websites, and systems and could be difficult to find. For example,

> "Searching through the websites there's all these different websites and it can be difficult. Keeping up with that, like I, you know, I was trying to read on various things and you know, having one concise place for everything would be nice and maybe that's out there now. I don't realize it."
>
> USGS data manager

Discussion

Attitudes toward data management planning and DMPs revealed in the survey responses are typical of those seen in other similar studies (Bishoff and Johnston, 2015; Simms and others, 2017; Smale and others, 2018; Hudson-Vitale and Moulaison-Sandy, 2019; Jones and others, 2020; Smale and others, 2020; Tenopir and others, 2020; and Australian Academy of Science, 2021). Additionally, the 30 percent or greater response rate from each survey group indicates our surveys provided a representative sample of USGS staff from across the Bureau (fig. 1). The survey responses revealed there are numerous opportunities for the Bureau to improve guidance and clarity regarding DMP purpose(s) and benefits. There also is a clear need for enhanced human infrastructure, training, tools, and resources to support data management and DMPs at the USGS. The USGS may also need to develop a strategy, other than through DMPs, for teaching and encouraging good data management practices. These surveys were an opportunity for USGS staff to provide feedback on their experiences. The surveys also revealed the need to support more regular evaluations, cross-disciplinary communication, and training on research data management (RDM) and

DMP development and integration in the context of USGS policy, FSP requirements, and overall Bureau expectations. For example,

> "USGS has a healthy obsession with data. This [comprehensive approach] is often difficult to convey to partners, especially because it adds time, cost, and complexity to interpretive and data programs. Efficient collection, QA/QC [quality assurance and quality control], and documentation are important, as is readily accessible storage and easy retrieval. Plans are a good step, but the data products need to be analyzed to refine the plans for continuous improvement."
>
> USGS Center Director

In the sections below, we discuss the survey results in the context of other literature. We also offer recommendations that the Science Data Management Branch, with the assistance of the USGS Community for Data Integration and the USGS Associate Chief Data Officer, can consider implementing to improve the value and usage of DMPs within the Bureau.

Data Management Plan Purpose and Benefits

Smale and others (2018) describe four purposes for DMPs: (1) to meet funder requirements for data sharing, (2) to gather institutional business intelligence, (3) to educate researchers or change behavior, and (4) to help researchers with project management planning. They argue that organizations often try to make DMPs serve all these purposes, which is nontrivial and likely an impossible task. Results from these surveys have demonstrated that the purpose for DMPs at the USGS is not clear and has been inconsistent in how it has been communicated. The USGS Survey Manual chapter 502.6 (USGS, 2017a) does not explicitly state the purpose for the DMP. The USGS Public Access Plan (USGS, 2016, p. 13) describes what should be in the DMP but also does

not explicitly state the purpose of DMPs. It states, "Prior to initiating research, intra- or extramural, approved plans must identify appropriate methods for digital data management, data release, and appropriate preservation in accordance with the USGS Records Disposition Schedules. The plans must also address making data available in appropriate long-term repositories (refer to section 8.1.6 [in the USGS Public Access Plan]) and stress the importance of nonproprietary, open formats for improved accessibility." Given that the Public Access Plan was developed in response to Federal requirements to increase public access to results from Federally funded research, it may be implied that the purpose of DMPs for intramural and extramural researchers is to ensure compliance with data management requirements, particularly data release and preservation.

The purpose of DMPs is not well-communicated across the Bureau. Many researchers in our survey were unsure about DMP requirements or the purpose and benefits of DMPs and were mainly motivated to create them because they were required by USGS policy or their Center Director (table 3). Data managers and Center Directors, on the other hand, believe that DMPs help researchers define how they will manage and track data throughout their projects, whereas a little less than one-half of the surveyed researchers stated they are motivated to create DMPs because they help them define how they will manage their data throughout the lifecycle of the project. Although this may be an added benefit of DMPs, perhaps it is not or should not be the main purpose.

There does seem to be agreement among these three groups that DMPs do not help researchers learn about new data management practices and requirements. For the most part, USGS researchers believe data management is important, and they do manage their data, but they may not document their plan or at least document it in the way USGS requires. Completing a DMP is not necessarily an educational experience (Smale and others, 2018) and the construction of DMPs should not be used as a data management training proxy. USGS will need to develop a different strategy for promoting leading data management practices and providing education on new requirements. For example, mandatory training on USGS FSP requirements rather than training on DMP development would help researchers know in advance that a USGS data release may be required. This awareness would help resolve some of the delays identified by survey respondents when they found themselves unprepared and lacking resources at the end of their project for comfortably accommodating this requirement.

Data management plans are variously described in existing literature as providing a written record of the data lifecycle within a project including, but not limited to the following: data collection and acquisition, processing, organization and storage (including data-related financial and IT-related requirements), documentation (metadata), quality assurance, access rights, data sharing, publication or release, and archiving and preserving (for example, Jones, 2011; Bishoff and Johnston, 2015; Hudson-Vitale and Moulaison-Sandy, 2019; Smale and others, 2020). However, Smale and others (2018) questioned

the view that good DMPs need to address the entire data lifecycle of a project. They also questioned the perceived or assumed benefits of DMP use (as distinct from the purpose[s] of DMPs). They suggested that rather than DMPs, organizations might consider requiring data sharing plans, or DSPs, which focus on describing compliance with data sharing policies. In addition to data sharing, there are various USGS policies that researchers are expected to adhere to, such as for records management (USGS, 2019a) and the Paperwork Reduction Act (USGS, 2019b). Focusing DMPs strictly on data sharing policies may be too narrow for USGS; however, narrowing the focus of DMPs to describe how project teams will meet specific USGS data-related policy requirements could reduce the amount of effort, time, and content they require.

The amount of time that it takes to create a DMP was listed as a major challenge by both researchers and data managers in our surveys. Implementation of more narrowly focused USGS DMPs may also make DMPs easier for data managers and Center Directors to evaluate. Researchers, data managers, and Center Directors all identified the most useful DMP section or element as the one that describes plans for data storage and backup, followed by data preservation and disposition plans and data release and publication (figs. 20–22). In free-text responses, survey respondents reported DMPs are most useful for ensuring FSP requirements are met, such as describing plans for data release and archiving. Some researchers and data managers also noted DMPs as being useful for aspects of their project planning such as budgeting time, money, and staff; however, these uses seemed to be an added benefit that some, but not most, researchers and data managers realize with DMPs.

Recommendations—

- Update USGS DMP guidance to narrow and explicitly state the intended purpose of DMPs to ensure USGS researchers plan to meet USGS and Federal policies related to data and information management.

- Develop communication strategies for informing USGS researchers, data managers, and Center Directors about the purpose and benefits of DMPs.

Human Infrastructure to Support Data Management Plans

Data management planning often focuses first on development of the workflows and cyberinfrastructure or machinery to facilitate it (Lowe, 1995). And although survey respondents suggested various machine-based solutions for improving DMP implementation at their Science Centers, their responses also highlighted the importance of an equally supported human infrastructure. Humans are necessary for feeding the machinery of data management (Lowe, 1995). Parr and McCarthy (2019) noted that data curators are critical links between end users (for example, which we interpret as both

researchers and final users of our data in the USGS context) and developers of data curation services and software (for example, the Science Data Management Branch of the USGS Science Analytics and Synthesis Program).

Researchers benefit from receiving help with data management, sharing, and archiving (Brandt, 2007; Smale and others, 2020; Tenopir and others, 2020). Data management planning and DMP creation should be a team effort that involves staff (beyond the principal investigator[s] and research team) with expertise in the data management, science, and technology domains (Peng and others, 2016, Miksa and others, 2019). Data manager involvement centers around ensuring DMP creation and review and approval. But where data managers are involved or help in DMP creation, researchers are significantly more likely to have a DMP associated with their most recent project and are also significantly more likely to view DMPs positively (as very or somewhat useful). Additionally, when there is a documented process for creating DMPs (that staff are aware of), researchers are significantly more likely to have a DMP associated with their most recent project. Similarly, researchers noted a need for dedicated staff who can provide support in creating DMPs and consultation on data management for specific data types.

The act of writing a DMP urges the writer to think about and try to anticipate their project's resource needs, including those necessary for generating and curating the products and data resulting from their project, and any possible challenges, risks, or dependencies this may entail. When researchers create a DMP alongside a data manager, the process may help researchers and data managers think about, discuss, and document decisions and solutions. As a result, project delays associated largely with a lack of awareness of what is expected of researchers throughout and especially at the end of the project may be avoided or at least mitigated. Having a data manager involved in creating or reviewing a DMP at the beginning of a project, even prior to its approval, would ensure there is a plan to meet USGS information management policy requirements. A review of a project's work plan and budget by a data manager might also reveal more time or resources are needed for the data release review and approval process. An analysis of the average time USGS data releases take from draft submission through the review and approval process to publication could also help researchers and data managers produce more accurate project schedules and budgets.

Researchers and Center Directors reported that some challenges in maintaining DMPs were the result of constantly evolving policies or projects. For events that cannot be anticipated, planning may not be precise, but involvement of a data manager in DMP creation at the beginning of a project, as well as regular reviews throughout the project lifecycle, could help address potential delays by increasing the overall data management skillset and domain awareness of the project team. In many of these scenarios, simply having a DMP may not be sufficient. Having someone with knowledge and expertise in USGS data management requirements review DMPs can successfully expose issues early in the project lifecycle. This concept of early identification of issues was one of the major themes from the surveys for how DMPs were useful to data managers. Regular reviews of DMPs with a data manager could also address the Center Directors' concerns that DMPs are not being followed.

The role of a data manager includes the ability to assess, label, and manage data quality (Lowe, 1995). They must be equipped to help authors meet Bureau publishing requirements while still maintaining the integrity, provenance, and correct characterization of information products as they are variously tagged, stored, indexed for search, and published for public consumption. Those in data management roles must also, to some extent, have good communication skills to effectively critique and deliver constructive feedback and instruction to policy makers, approving officials, researchers, and other stakeholders during the project and data management lifecycles.

Information technologists, for the sake of this study, were grouped in with data managers. This grouping is due, in part, to the responsibilities for data management and DMP-related roles (fig. 2) and activities at USGS often falling to staff with other, preexisting primary roles (for example, table 2) including IT professionals. Repository operators or IT professionals and those who provide other cyberinfrastructure and computing resources are often not involved nor even informed at the outset of a project about potential service demands (Miksa and others, 2019 and 2022). Simms and others (2017) also noted that repositories, or personnel involved in managing them, such as IT professionals and library scientists, rarely play an active role in the data management planning process. Information technologists should be more involved in USGS data management planning and DMP creation because they can provide guidance on cyberinfrastructure considerations and requirements (Miksa and others, 2019).

Recommendations—

- Empower data managers and information technologists in the beginning or planning stages of research projects (for example, prior to the commencement of data collection or acquisition) and in DMP development.

- Work with data managers to provide an analysis of the average time USGS data releases take from draft submission through review and approval and document this information on the USGS Data Management website.

Current State of Data Managers in the U.S. Geological Survey

A little under one-half of USGS Science Centers (43.9 percent) represented in our surveys had someone occupying a role(s) related to DMPs (fig. 2). Among the 37 different DOI Active Directory position titles of data managers with some role related to DMPs, only ten individuals held position titles that included the word "data" and four identified

as a "data management specialist" (table 2). This variety of titles is a consequence of the lag in formal recognition of data management responsibilities within the USGS. Only recently (as of 2019) has there been an official USGS data manager position description (PD) for use to hire staff into this role.

In July 2019, a subteam of the USGS Community for Data Integration Data Management Working Group drafted a PD for a Data Management Specialist, which was added to the PD database for access during hiring actions. This PD described, for the first time, the education, skills, roles, and responsibilities of a USGS data manager, paving the way for hiring managers to begin to build this new workforce sector at the USGS. It also served to document those skills, roles, and responsibilities of staff hired for other purposes, and in non-data-manager series, while fulfilling this additional role.

The new PD helps the USGS hire personnel to assist researchers with creating and managing their DMPs. However, the Bureau could benefit from a greater number and increased availability of full-time professional data managers to ensure full coverage across all Science Centers and programs. In the meantime, Science Centers with few resources could share staff or receive internal consulting support from other Science Centers' data managers to help researchers meet USGS FSP and data management requirements.

Recommendation—

- Work with Center Directors to help them understand the benefits of assigning a full-time data manager for every Science Center, preferably as a full-time member of the Science Center or minimally as a shared resource across a set of Science Centers.

Data Management and Policy Training for Researchers and Data Managers

If the USGS expects to increase data management staff and to have them take on a bigger role with respect to helping create DMPs, then we need to provide them with the training to be successful in this role. Some of the inconsistencies of DMP implementation across different USGS Science Centers and Mission Areas may be because of a lack of training and detailed guidance at the Bureau level. In many cases, budget and Federal hiring challenges have also necessitated the responsibility for data management, data management planning, and development of DMPs to be assigned to staff from a broad range of educational backgrounds, skillsets, and areas of expertise, with position titles (for example, see table 2) and positions within the Bureau not specifically or typically associated with career data managers.

We did not specifically include questions in our survey about whether data managers answer questions or help researchers write DMPs, but many voluntarily provided this information. One data manager interviewed wanted training so they could better answer questions from researchers regarding DMPs and what information should be contained in a DMP. These results provide evidence of a desire among data managers for more training and knowledge to better equip them for the role of teaching and assisting others in data management activities, DMP creation, and meeting Bureau data management requirements.

In turn, increased data management training for other staff by data managers may result in overall increased efficiency and adoption of data management planning. It may also result in an increase in quality and timeliness (or a reduction in publishing delays) of USGS data releases and related publications. However, it should also be noted that training in DMP creation or DMP completion does not translate automatically or directly into improved data management practices (Smale and others, 2020).

If the USGS agrees that data management is a role with a valuable skillset and deserving of credit, we should also not expect all researchers to be experts in data management and information sciences. U.S. Geological Survey researchers, although already well versed in an understanding of the scientific method, attention to detail, and effective documentation for reproducibility, have been asked to also increase their understanding of digital data management and the information sciences. These newer and rapidly evolving areas of expertise require an understanding of content, metadata, and other digital data standards. A basic knowledge of IT solutions for documenting, packaging, and formatting data for release in open-source file formats also is increasingly necessary among all involved in data management activities. The USGS has expected all these new skills to be gained while also continuing to maintain an awareness of policy and technological requirements of data repositories. More involvement from data managers and information technologists in project planning and throughout the course of a project could help reduce the amount of data management and information science expertise a researcher is expected to have. This involvement would likely reduce delays at the end of a project caused by researchers' lack of awareness of information management requirements.

Recommendation—

- Develop training materials to help data managers feel empowered to assist and train researchers in their Science Centers on data and information management policies and DMP creation.

Resource Needs for Researchers, Data Managers, and Center Directors

Alongside training, USGS researchers, data managers, and Center Directors would benefit from the development of additional resources to facilitate and support DMP creation, maintenance, and curation.

Templates and Tools for Creating Data Management Plans

The effort involved in creating and maintaining DMPs should not outweigh their usefulness. Poor-quality DMPs and a lack of adherence to following DMPs makes them ineffective and of limited benefit to researchers, institutions, and funding bodies (Smale and others, 2018; Miksa and others, 2022). Tools for DMP creation can help drive the creation of quality DMP content and fitness for purpose. Templates have a useful role in USGS DMP development and may also facilitate increased machine readability and integration of DMPs. However, poorly designed or highly restrictive templates may also discourage engagement in the DMP creation process (Smale and others, 2020).

Center Directors were asked how the implementation of DMPs at their Centers could be improved. More than one-fourth of their free-text responses indicated that a universal, well-thought-out, web-based tool and template would help with DMP implementation at their Center. One Center Director indicated that their Center is "pretty set" because they have a DMP tool that facilitates DMP creation, records management, and approval.

Researcher survey respondents saw value in DMP templates for improving the usefulness of DMPs, but they also observed difficulties in knowing how to respond to or fill in certain template sections. This uncertainty, in turn, can result in missing details and poor-quality DMPs (Hudson-Vitale and Moulaison-Sandy, 2019). USGS DMPs are also perceived by researchers as involving a lot of duplication of effort and information, often already captured in other internal project documentation. One solution might be to create standardized text at the Bureau level for specific DMP sections or even a Bureau-level DMP. Some examples of existing USGS data management planning language that may be useful for standardization include language already extensively developed as part of the USGS SM and FSP; policy language regarding the management of personally identifiable information in official information products (for example, USGS, 2022a and 2022b); language about the documentation of data backup and preservation plans in USGS staff exit surveys (for example, USGS, 2014 and 2022c); and USGS standard disclaimer statements (OSQI, 2019). Data management plans created at the level of individual programs and research projects could then have the option to reference or point to standard text or sections in the Bureau-level DMP rather than duplicating the same content or some variation of it.

Researchers predominantly requested more flexibility and the ability to customize DMPs, contrasting with data managers who preferred to see more standardization in DMPs. And although data managers indicated that more standardization could improve the usefulness of DMPs, they also recognized standardization may be difficult due to the variation in projects. This contrast may be explained by the different needs of these two DMP stakeholder groups. Researchers seeking to create comprehensive and detailed DMPs that accurately document their individual project data management activities and serve as a useful reference throughout their project, understandably need flexibility to do so. But the increased uniqueness of DMPs and their non-machine-readable narrative format makes them difficult and very time consuming for data managers, Center Directors, and BAOs to evaluate and approve them, which is likely also a factor in USGS DMPs being perceived as only being evaluated for their presence rather than the quality of their contents.

If we agree that the purpose of the USGS DMP is to ensure compliance with USGS information management policies, then DMPs should be highly standardized. The flexibility requested by researchers to document their other project lifecycle data management activities could be accommodated through mechanisms other than DMPs. Because that information would not be required by USGS policy for Center Directors or BAOs to review, there would not need to be standardization. Smale and others (2018) note that research groups managing complex projects with complex data will treat data "as an intrinsic and underpinning component of the research itself." Much of the information related to data acquisition, organization, processing, and analysis should be allowed to be captured in a way that is useful to the project team and not necessarily part of a formal DMP template (Smale and others, 2018). In the USGS, data managers can be available to help project teams think through and answer questions on various aspects of data management but not require them to document this information in a specific or standardized format.

Regardless of what the final templates look like (flexible or standardized), USGS stakeholders (researchers, data managers, information technologists, and Center Directors) need to be involved in the development of DMP templates and tools. Hudson-Vitale and Moulaison-Sandy (2019) noted in their review that research on DMPs or their use finds that both DMPs and the process(es) used to create them are "* * * largely ineffective." (p. 323). This finding is likely because all stakeholders were not involved in the creation of processes and tools.

Recommendations—

- Develop standardized text at the Bureau level that addresses how to comply with certain Federal and USGS information management policy requirements for easy inclusion in DMPs.

- Work with stakeholders, such as researchers, data managers, information technologists, Center Directors, and BAOs to update templates and tools. By updating these resources, researchers and data managers will have the ability to more easily incorporate standardized text into their DMPs in a machine-readable format.

- Work with stakeholders to update USGS DMP templates to assist researchers in documenting how they intend to meet all relevant data and information management policies. Other information related to data acquisition, organization, processing, and analysis should be allowed to be captured outside of the DMP template to ensure that it is useful to the project team.

Example Processes and Workflows for Creating Data Management Plans

In addition to using DMP templates, data managers and Center Directors indicated processes and workflows need to be established at the Center level and be supported by Center Directors and accepted by researchers. Survey respondents noted better integration of DMPs into existing Center-level project workflows could facilitate conversations around project and data management planning among data managers, information technologists, and researchers. Based on survey responses, the workflows also need to include and facilitate, rather than impede, periodic reviews and updates of DMPs and data management activities throughout the project lifecycle. DMP reviews and updates may also be achieved through increased DMP automation, such as pulling information in from relevant systems and automated notifications for reviewing and updating the DMP.

Most researchers (50.5 percent) reported that their Center has a documented process for creating DMPs. However, a large proportion of researchers were not sure (41.2 percent) or reported that their Center does not have a documented process for creating DMPs (8.2 percent) (fig. 7). This uncertainty speaks to an opportunity for the Bureau to encourage and provide guidance and support to USGS Science Centers on instituting DMP creation processes and educating staff about them, or to consider establishing a core set of DMP creation process steps that could be adopted Bureau wide while still allowing flexibility at the Center level.

Smale and others (2020) suggested DMP mandates are likely ineffective at creating cultural change but concluded that the questions asked by DMPs "may play some part in a cultural shift in research towards consideration of data life-cycle issues" (p. 23). The researcher and data manager interviews revealed that the presence of both a Center Director and a data manager who supported and encouraged good data management practices appeared to have a positive effect on Center data management culture. One-half of researchers in the survey (50.3 percent) noted they were motivated to create DMPs because it is required by the Center Director. Researchers and data managers take their lead from their Center Director. Center Directors play a critical role in hiring data managers and providing resources and guidance for data managers so they can establish processes, develop training and tools, and be involved in projects at specific points during the project lifecycle.

Recommendations—

- Provide examples of DMP creation procedures from Science Centers with existing processes in place or develop DMP creation processes that can be adopted Bureau wide.

- Encourage Center Directors and data managers to emphasize the importance of data management and DMP creation to researchers at their Centers.

Example Processes and Checklists for Reviewing and Approving Data Management Plans

Only 33 of the 80 (about 40 percent) USGS Science Centers represented in all survey responses reported having a documented process for DMP review and approval. The perceived limited effect or lack of consequences associated with not meeting DMP creation requirements also is an impediment to their adoption. For example, of the researchers in our survey who had never participated in DMP creation, the majority (62.2 percent) reported they had never been required to create one. Mischo and others (2014), Bishoff and Johnston (2015), and Mannheimer (2018) found the inclusion of or degree of completeness of DMPs in project proposals made no difference in the funding of projects or success of grant proposals. With little guidance on how the contents of project DMPs should be evaluated, the treatment by approving authorities has, in some cases, become little more than an act of checking a box for whether a DMP was included in a project proposal. The sentiments of some USGS researchers during our survey echoed this common perception of DMPs reported in other studies as nothing more than an administrative burden and "box checking" requirement offering little to no benefit to the project or to RDM (Lowe, 1995; Simms and others, 2017; Hudson-Vitale and Moulaison-Sandy, 2019; Miksa and others, 2019; Miksa and others, 2022; Smale and others, 2020).

The use of DMPs as compliance tools is only as effective as their monitoring and subsequent remediation (Smale and others, 2020). The USGS Public Access Plan (USGS, 2016, p. 14) states that "FSP policies ensure compliance" with approved data management plans, and that compliance "will continue to be ensured through progress reporting as required in the funding agreement and the Financial Assistance Monitoring Protocol used by USGS pursuant to 2 CFR 200.205(c)(3). Starting in January 2016, USGS will be required to report such recipients to Federal Awardee Performance and Integrity Information System (FAPIIS) as required by 2 CFR 200.212. Funds are withheld if an awardee is in noncompliance* * *." However, it is unclear if these protocols and systems are being used. Additionally, section 9.3 "Evaluation of Data Management Plans" in the Public Access Plan states, "USGS FSP policy describes the process for evaluating data management plans in the overall research review process" and "these processes are elucidated at the USGS Data Management website where detailed guidance is provided in the form of explanatory text and checklists to ensure appropriate evaluation of the merits of submitted data management plans by research proposal reviewers" (page 13). However, these instructional resources are not readily apparent on the referenced website, and if they do exist, they may not be easily found by USGS staff. Providing guidance, defined criteria, tools, and training on how to evaluate USGS DMPs effectively and efficiently may help Center Directors and BAOs provide useful feedback to DMP creators. Receiving feedback and an increased understanding of how DMPs will be evaluated may

help researchers and others develop better-quality DMPs, as well as increase the perception that DMPs are read and used by others.

> **Recommendations—**
>
> • Provide examples of DMP review and approval procedures from Science Centers with existing processes or develop DMP review and approval processes that can be adopted Bureau wide.
>
> • Define criteria and develop checklists for assessing and approving DMPs. Make these checklists available and easily discoverable through the USGS Data Management website, similar to the criteria and checklists already available for metadata and data review.

Data Management Plan Accessibility and Storage

Bureau Approving Officials expressed the need for DMPs to be in a location that is accessible to them to support their review and evaluation responsibilities. Researcher, data manager, and Center Director survey responses, as well as some interviewees, also indicated the need for a Center-level, shared location for DMP storage and access, and guidance about its use. This infrastructure solution would need to be accessible by all stakeholders, including BAOs, it would need a complementary communication and awareness-raising campaign about its existence (fig. 12), and it may also need to accommodate DMP maintenance as living documents.

The FAIR guiding principles were developed to clarify what good data management means (Wilkinson and others, 2016). The data lifecycle implementation choices made and documented in a DMP can affect the degree of conformity with the FAIR guiding principles of a project's products (Wilkinson and others, 2016). If all research objects should follow FAIR principles, there is increasing recognition that in addition to data, DMPs should also partially or completely follow FAIR principles (Jones and others, 2020; Miksa and others, 2019; Simms and others, 2017; Wilkinson and others, 2016). Although the Bureau provides guidance on officially recognized solutions for data publishing and storage in trusted digital repositories (for example, Hutchison and others, 2021), USGS lacks the same type of guidance for DMPs. Also, whether maintained as static or living documents, the accessibility of USGS DMPs to the stakeholders identified in this study is limited temporally (for example, to specific periods of the project lifecycle such as proposal evaluation or data publication) or physically (for example, storage on individuals' computers, or on limited-membership network locations). Simms and others (2017) suggested that increasing the "FAIRness" of DMPs, even in unstructured formats lacking persistent identifiers or versioning, can be valuable and incentivize the creation and maintenance of good DMPs.

Survey respondents noted the value of being able to see and emulate other examples of good DMPs. This notion indicates another reason for improving DMP access at least within the Bureau. There was strong support from survey respondents for increasing the accessibility of DMPs internal to the Bureau, but equally strong resistance to, and disagreement with, the idea of publishing or making USGS DMPs publicly available. Some survey respondents expressed hesitation in documenting their DMPs in writing over concerns that they might be penalized for not following them.

Prior to the survey, the survey administrator's exposure to USGS DMPs was limited to a few examples on the USGS Data Management website and anecdotes from other data managers. However, the volume of survey repondents indicates a higher level of DMP adoption by USGS stakeholders than survey administrators expected (fig. 4). Survey administrators' lack of evidence or awareness of DMPs may be explained, in part, by the lack of transparency and accessibility of USGS DMP storage and management mechanisms. The discrepancy between researchers (32.4 percent) and data managers (13.5 percent) reporting on availability of DMPs as "upon request from a project team member's computer" (fig. 11) may be because data manager responses represent Centers where data managers have a role related to DMPs. This result may indicate that as data managers get involved with DMP development, the number of DMPs that are only available through request from a team member's computer may decrease. Moreover, DMP accessibility becomes more aligned with FAIR principles when data managers are involved in the process.

> **Recommendation—**
>
> • Encourage every Science Center to implement a single shared location for their DMP storage and access that accommodates maintaining DMPs as living documents, allows for persistent links to the most current version of a DMP, and enables read-only access to be given to anyone in the USGS, especially Bureau Approving Officials.

Data Management Plans as Static versus Living Documents

The maintenance of DMPs as static documents can be difficult and cumbersome (Hudson-Vitale and Moulaison-Sandy, 2019; Miksa and others, 2019 and 2022; Simms and others, 2017; Smale and others, 2020). Survey respondents who said DMPs are useful also expressed that one of the biggest challenges is maintaining DMPs, particularly in a static document format (fig. 14). A lack of usefulness and user-friendliness in processes and workflows can discourage the adoption, use, and user participation in DMPs. Although Mannheimer (2018) found that principal investigators did not use or reference their DMPs as guiding documents during their research projects, at least one data manager in this study noted that DMPs might be more useful if principal investigators had more incentive

to complete and maintain them throughout the project life-cycle. Hudson-Vitale and Moulaison-Sandy (2019) assert that an inability to edit DMPs may be one reason why some of the early research into DMP adoption and compliance has shown them to be ineffective. Good document management and control skills and having the time to perform these activities are critical to the successful creation and maintenance of DMPs, especially if they are multiauthored static documents. Researchers need the ability to update their DMPs throughout the project lifecycle to accommodate changes to their project or USGS policies.

Smale and others (2020) observed a gap between the ideal versus actual data management practices of researchers. Many researchers' survey responses described their Centers as having a policy for updating DMPs and that DMPs are much more useful as living documents, but it became evident in subsequent interviews of survey participants that the actual updating of DMPs was often not happening. Good intentions to maintain DMPs as living documents are evident and formally recognized in Center policies, but perhaps the necessary time, tools, workflows, and motivations are unavailable. Although respondents indicated maintaining DMPs as living documents was a move in the right direction, some also recognized that better mechanisms than those typically used in basic document editing are needed to facilitate efficiencies and benefits of living DMPs. To accommodate changes in RDM that occur during research projects and increase the usefulness of DMPs (as measured by how accurately they represent the data management activities of a project over time), we recommend that USGS DMPs be maintained as living documents if not as partial to fully machine-actionable data management plans (maDMPs; Jones and others, 2020; Simms and others, 2017; Miksa and others, 2019; Miksa and others, 2022). Machine-actionable data management plans are those that "improve the experience for all involved by exchanging information across research tools and systems and embedding DMPs in existing workflows" (Miksa and others, 2019).

Recommendations—

- Provide example workflows on the USGS Data Management website for how researchers and data managers can incorporate DMP updates into their project lifecycle.

- Develop tools and workflows to facilitate the updating and version control of DMPs throughout the project lifecycle.

Conclusions

A major component of successful workflows is ensuring buy-in from all stakeholders. Data management plans (DMPs) are being used to serve many purposes without a clearly defined audience or reason. In addition to documenting the numerous data management planning elements in a project's

data lifecycle, meeting the needs of a broad group of stakeholders may be beyond the abilities of a single, static DMP. If the U.S. Geological Survey (USGS) were to establish the main purpose of DMPs as ensuring that project teams will meet Federal, departmental, and Bureau information management policy requirements, it will be easier to establish effective templates and tools, develop appropriate processes, and communicate the benefits of DMPs to all stakeholders.

Data management plan templates and tools can be improved to include guidance on all potentially relevant information management policies and standardized text or recommendations on how project teams can meet these policies. With a focus on meeting policy requirements, DMPs could be simplified, thus reducing the time and effort to complete them, a barrier to usage and updates. To ensure these templates and tools are useful, a representation of all stakeholders needs to be involved in the development process.

Data managers and information technologists are a critical component to ensuring the success of projects in meeting requirements of information management policies and the development of DMPs. These specialists need to have clear direction and advocacy from their Center Directors. Center Director support can ensure project teams develop an effective DMP, discuss data and information management needs, establish regular progress reviews for project data management activities, and help ensure projects are on track to meet information management policy requirements. Data managers and information technologists in the Bureau have, until recently, represented a somewhat grassroots and perhaps under-supported effort to serve the Bureau's requirements for meeting Federal findable, accessible, interoperable, and reusable (FAIR) data documentation, publishing, and access requirements. They also serve individual researchers' needs for assistance in navigating USGS policy and the digital research data management world. These USGS staff are a small but growing group of advocates for encouraging adoption of good data management practices and ensuring compliance with policies, such as the relatively new DMP requirement. Survey administrators believe providing credit and promotion opportunities would help with the attraction and retention of data management specialists.

Narrowing the purpose and including data managers in the development of DMPs will enable the USGS to establish and communicate the benefit of DMPs more clearly and broadly to all stakeholders. Well-defined DMPs will allow researchers to feel confident they are meeting USGS and other Federal requirements and help ensure their publications will not be delayed due to unmet requirements. Data managers and information technologists will be aware of project needs from the beginning and be more prepared to support project teams with fewer "emergencies" to meet policy at the end of the project. Center Directors will feel more confident that resources will be efficiently spent within projects with less duplication of effort due to project teams needing to repeat or revise work that is out of compliance with Federal or USGS policies. Bureau Approving Officials, with access to the most

recent DMPs, will be able to check more easily that relevant data management policies were met when approving a publication.

In the future, USGS could explore opportunities for implementing machine-actionable data management plans (maDMPs). The term "machine-actionable" refers to the principle that machines should be able to act on digital objects (Miksa and others, 2022), and this principle also is associated with the FAIR principles (Miksa and others, 2019 and 2022; Wilkinson and others, 2016). Structured, machine-actionable USGS DMPs present a possible solution for realizing the full potential of DMPs, making them compliant with FAIR principles, and improving the overall user experience for all USGS stakeholders involved in data management planning. In recent years, there has been encouraging progress in developing criteria and solutions for implementing DMPs as living documents or maDMPs (for example, Jones and others, 2020; Simms and others, 2017; Miksa and others, 2019 and 2022). Interest in implementing maDMPs and increasing integration of USGS DMPs with existing cyberinfrastructure as part of USGS Fundamental Science Practices and data publishing workflows was also evident in survey and interviewee responses. Even basic improvements, such as automated population of user-selected DMP sections with standardized information (for example, information about people, organizations, and budgets) sourced from other USGS information management systems and referenced by persistent identifiers of various types, could streamline USGS DMP creation and maintenance. The content of maDMPs could also trigger certain events or notifications at appropriate points throughout the data lifecycle (Miksa and others, 2019 and 2021; Simms et al, 2017) based on user inputs. For example, if the project team notes a data release will be needed and identifies the intended repository, then the repository managers could be notified. Or, if the project team indicates the Paperwork Reduction Act is relevant to their project in their DMP, then the maDMP could trigger the initiation of that process.

Finally, given that survey groups indicated a comparatively high awareness of the USGS Data Management website as a resource (table 5), it would be a logical vehicle for providing additional information, guidance, and training materials about DMPs. However, there was also evidence that USGS staff awareness of DMP-related resources could be higher if they were provided in more places and promoted more often.

The final recommendations for consideration based on these surveys and interviews are as follows:

1. Update USGS DMP guidance to narrow and explicitly state the intended purpose of DMPs to ensure USGS researchers plan to meet USGS and Federal policies related to data and information management.

2. Develop communication strategies for informing USGS researchers, data managers, and Center Directors about the purpose and benefits of DMPs.

3. Empower data managers and information technologists in the beginning or planning stages of research projects (for example, prior to the commencement of data collection or acquisition) and in DMP development. This expectation should be clearly and specifically communicated in policy and training materials.

4. Work with data managers to provide an analysis of the average time that USGS data releases take from draft submission through review and approval, and document this information on the USGS Data Management website.

5. Work with Center Directors to help them understand the benefits of assigning a full-time data manager for every Science Center, preferably as a full-time member of the Science Center or minimally as a shared resource across a set of Science Centers.

6. Develop training materials to help data managers feel empowered to assist and train researchers in their Science Centers on data and information management policies and DMP creation.

7. Develop standardized text at the Bureau level that addresses how to comply with certain Federal and USGS information management policies for easy inclusion in DMPs.

8. Work with stakeholders, such as researchers, data managers, information technologists, Center Directors, and Bureau Approving Officials to update templates and tools. By updating these resources, researchers and data managers will have the ability to more easily incorporate standardized text into their DMPs in a machine-readable format.

9. Work with stakeholders to update USGS DMP templates to assist researchers in documenting how they intend to meet all relevant data and information management policies. Other information related to data acquisition, organization, processing, and analysis should be allowed to be captured outside of the DMP template to ensure that it is useful to the project team.

10. Provide examples of DMP creation procedures from Science Centers with existing processes or develop DMP creation processes that can be adopted Bureau wide.

11. Encourage Center Directors and data managers to emphasize the importance of data management and DMP creation to researchers at their Centers.

12. Provide examples of DMP review and approval procedures from Science Centers with existing processes or develop DMP review and approval processes that can be adopted Bureau wide.

13. Define criteria and develop checklists for assessing and approving DMPs and make the checklists available and easily discoverable through the USGS Data Management website, similar to the criteria and checklists already available for metadata and data review.

14. Encourage every Science Center to implement a single shared location for their DMP storage and access that accommodates maintaining DMPs as living documents, allows for persistent links to the most current version of a DMP, and enables read-only access to be given to anyone in the USGS, especially Bureau Approving Officials.

15. Provide example workflows on the USGS Data Management website for how researchers and data managers can incorporate DMP updates into their project lifecycle.

16. Develop tools and workflows to facilitate the updating and version control of DMPs throughout the project lifecycle.

The survey administrators will work with the USGS data management community to prioritize these recommendations. They will also help to develop a plan for implementing the highest priority recommendations to increase the value and usage of DMPs within the USGS.

Acknowledgments

References Cited

Australian Academy of Science, 2021, Advancing data-intensive research in Australia: Canberra, Australia, Australia Academy of Science, 76 p., accessed October 19, 2021, at https://www.science.org.au/supporting-science/science-policy-and-analysis/reports-and-publications/advancing-data-intensive-research-australia.

Bishoff, C., and Johnston, L., 2015, Approaches to data sharing—An analysis of NSF data management plans from a large research university: Journal of Librarianship and Scholarly Communication, v. 3, no. 2, p. 1231, accessed June 10, 2022, at https://doi.org/10.7710/2162-3309.1231.

Brandt, D.S., 2007, Librarians as partners in e-research—Purdue University Libraries promote collaboration: College & Research Libraries News, v. 68, no. 6, p. 365–396, accessed March 2, 2022, at https://doi.org/10.5860/crln.68.6.7818.

Hudson-Vitale, C., and Moulaison-Sandy, H., 2019, Data management plans—A Review: DESIDOC Journal of Library and Information Technology, v. 39, no. 6, p. 322–328, accessed March 2, 2022, at https://doi.org/10.14429/djlit.39.06.15086.

Hutchison, V.B., Norkin, T., Langseth, M.L., Ignizio, D.A., Zolly, L.S., McClees-Funinan, R., and Liford, A., 2021, Leveraging existing technology—Developing a trusted digital repository for the U.S. Geological Survey: International Journal of Digital Curation, v. 16, no. 1, p. 23, accessed March 2, 2022, at https://doi.org/10.2218/ijdc.v16i1.741.

Jones, S., 2011, How to develop a data management and sharing plan: Edinburgh, Scotland, Digital Curation Centre, accessed June 10, 2022, at https://www.dcc.ac.uk/guidance/how-guides/develop-data-plan.

Jones, S., Pergl, R., Hooft, R., Miksa, T., Samors, R., Ungvari, J., Davis, R.I., and Lee, T., 2020, Data management planning—How requirements and solutions are beginning to converge: Data Intelligence, v. 2, no. 1–2, p. 208–219, accessed March 18, 2022, at https://doi.org/10.1162/dint_a_00043.

Joshi, A., Kale, S., Chandel, S., and Pal, D., 2015, Likert scale—Explored and explained: British Journal of Applied Science and Technology, v. 7, no. 4, p. 396–403, accessed May 12, 2022, at https://doi.org/10.9734/BJAST/2015/14975.

Langseth, M.L., Donovan, G.C., Liford, A.N., and Sellers, E.A., 2023, U.S. Geological Survey 2021 data management planning survey results and analyses: U.S. Geological Survey data release, available at https://doi.org/10.5066/P91WKCA3.

Lowe, D.J., 1995, The geological data manager—An expanding role to fill a rapidly growing need: Geological Society Special Publication, v. 97, p. 81–90, accessed March 2, 2022, at https://doi.org/10.1144/GSL.SP.1995.097.01.10.

Mannheimer, S., 2018, Toward a better data management plan—The impact of DMPs on grant funded research practices: Journal of Escience Librarianship, v. 7, no. 3, p. e1155, accessed March 15, 2022, at https://doi.org/10.7191/jeslib.2018.1155.

Mischo, W., Schlembach, M., and O'Donnell, M., 2014, An analysis of data management plans in University of Illinois National Science Foundation grant proposals: Journal of Escience Librarianship, v. 3, no. 1, accessed March 15, 2022, at https://doi.org/10.7191/jeslib.2014.1060.

Miksa, T., Simms, S., Mietchen, D., and Jones, S., 2019, Ten principles for machine-actionable data management plans: PLoS Computational Biology, v. 15, no. 3, p. e1006750, accessed March 4, 2022, at https://doi.org/10.1371/journal.pcbi.1006750.

Miksa, T., Walk, P., Neish, P., Oblasser, S., Murray, H., Renner, T., Jacquemot-Perbal, M.-C., Cardoso, J., Kvamme, T., Praetzellis, M., Suchánek, M., Hooft, R., Faure, B., Moa, H., Hasan, A., and Jones, S., 2021, Application profile for machine-actionable data management plans: Data Science Journal, v. 20, no. 1, p. 32, accessed March 4, 2022, at https://doi.org/10.5334/dsj-2021-032.

Miksa, T., Oblasser, S., and Rauber, A., 2022, Automating research data management ssing machine-actionable data management plans: ACM Transactions on Management Information Systems, v. 13, no. 2, p. 1–22, accessed March 4, 2022, at https://doi.org/10.1145/3490396.

Office of Science and Technology Policy, 2013, Memorandum for the Heads of Executive Departments and Agencies—Increasing access to the results of federally funded scientific research: Washington, D.C., Executive Office of the President, 6 p., accessed May 11, 2022, at https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf.

Office of Science Quality and Integrity (OSQI), 2019, Fundamental Science Practices (FSP) guidance on disclaimer statements allowed in USGS science information products: U.S. Geological Survey website, accessed March 31, 2022, at https://www.usgs.gov/about/organization/science-support/office-science-quality-and-integrity/fundamental-science-5.

Parr, C., and McCarthy, S., 2019, Building capacity for data science with help from our friends: Research Library Issues, (298): p. 28–40, accessed March 15, 2022, at https://doi.org/10.29242/rli.298.4.

Peng, G., Ritchey, N.A., Casey, K.S., Kearns, E.J., Prevette, J.L., Saunders, D., Jones, P., Maycock, T., and Ansari, S., 2016, Scientific stewardship in the open data and big data era—Roles and responsibilities of stewards and other major product stakeholders: D-Lib Magazine : the Magazine of the Digital Library Forum, v. 22, no. 5/6, accessed March 2, 2022, at https://doi.org/10.1045/may2016-peng.

Simms, S., Jones, S., Mietchen, D., and Miksa, T., 2017, Machine-actionable data management plans (maDMPs): Research Ideas and Outcomes, v. 3, p. e13086, accessed March 2, 2022, at https://doi.org/10.3897/rio.3.e13086.

Smale, N., Unsworth, K., Denyer, G., Magatova, E., and Barr, D., 2018, The history, advocacy and efficacy of data management plans: bioRxiv, 30 p., accessed October 27, 2020, at https://doi.org/10.1101/443499.

Smale, N.A., Unsworth, K., Denyer, G., Magatova, E., and Barr, D., 2020, A review of the history, advocacy and efficacy of data management plans: International Journal of Digital Curation, v. 15, no. 1, p. 30, accessed March 2, 2022, at https://doi.org/10.2218/ijdc.v15i1.525.

Tenopir, C., Rice, N.M., Allard, S., Baird, L., Borycz, J., Christian, L., Grant, B., Olendorf, R., and Sandusky, R.J., 2020, Data sharing, management, use, and reuse—Practices and perceptions of scientists worldwide: PLoS One, v. 15, no. 3, p. e0229003, accessed March 2, 2022, at https://doi.org/10.1371/journal.pone.0229003.

U.S. Geological Survey (USGS), 2011, Fundamental science practices—Planning and conducting science research: U.S. Geological Survey Manual, chap. 502.2, accessed February 1, 2022, at https://www.usgs.gov/survey-manual/5022-fundamental-science-practices-planning-and-conducting-data-collection-and.

U.S. Geological Survey (USGS), 2014, USGS science data exit survey form (ver. 1, November 2014): U.S. Geological Survey website, accessed March 31, 2022, at https://www.usgs.gov/media/files/usgs-science-data-exit-survey-form.

U.S. Geological Survey (USGS), 2015, Scientific data management: U.S. Geological Survey Instructional Memorandum IM OSQI 2015–01. [Superseded in 2017 by USGS Manual chapter 502.6, which is available at https://www.usgs.gov/survey-manual/5026-fundamental-science-practices-scientific-data-management.]

U.S. Geological Survey (USGS), 2016, Public access to results of federally funded research at the U.S. Geological Survey—Scholarly publications and digital data: U.S. Geological Survey, 18 p., 2 app., accessed March 18, 2022, at https://www.usgs.gov/media/files/public-access-results-federally-funded-research-us-geological-survey-scholarly.

U.S. Geological Survey (USGS), 2017a, Fundamental science practices—Scientific data management, U.S. Geological Survey Manual, chap. 502.6, accessed February 1, 2022, at https://www.usgs.gov/survey-manual/5026-fundamental-science-practices-scientific-data-management.

U.S. Geological Survey (USGS), 2017b, Fundamental science practices—Preservation requirements for digital scientific data, U.S. Geological Survey Manual, chap. 502.9, accessed February 1, 2022, at https://www.usgs.gov/survey-manual/5029-fundamental-science-practices-preservation-requirements-digital-scientific-data.

U.S. Geological Survey (USGS), 2019a, Records management roles and responsibilities, U.S. Geological Survey Manual, chap. 431.1,– accessed April 5, 2022, at https://www.usgs.gov/survey-manual/4311-records-management-roles-and-responsibilities.

U.S. Geological Survey (USGS), 2019b, Information collection requirements: U.S. Geological Survey Manual, chap. 431.10, accessed April 5, 2022, at https://www.usgs.gov/survey-manual/43110-information-collection-requirements.

U.S. Geological Survey (USGS), 2021, Data management plans: U.S. Geological Survey website, accessed February 1, 2022, at https://www.usgs.gov/data-management/data-management-plans.

U.S. Geological Survey (USGS), 2022a, Privacy policies: U.S. Geological Survey website, accessed March 31, 2022, at https://www.usgs.gov/office-of-the-director/privacy-policies.

U.S. Geological Survey (USGS), 2022b, E.6 Software: U.S. Geological Survey Extended Guidance and Specific Products, chap. E.6.5, accessed March 31, 2022, at https://www.usgs.gov/office-of-science-quality-and-integrity/e6-software#6.5.

U.S. Geological Survey (USGS), 2022c, USGS science data exit survey: U.S. Geological Survey website, accessed March 31, 2022, at https://www.usgs.gov/data-management/usgs-science-data-exit-survey.

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Bloomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., 't Hoen, P.A.C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., and Mons, B., 2016, The FAIR guiding principles for scientific data management and stewardship: Scientific Data, v. 3, article 160018, 9 p., accessed March 18, 2022, at https://doi.org/10.1038/sdata.2016.18.

Appendixes

The appendixes in this report are available externally at https://pubs.usgs.gov/publication/ofr20231069. The original survey documents are provided for context and have not been modified.

Appendix 1.   Data Management Planning Questionnaire for Researchers

This appendix, available at https://pubs.usgs.gov/publication/ofr20231069, includes a copy of the survey administered to researchers. The survey consisted of multiple choice, multiple selection, Likert scale, and free-text questions.

Appendix 2.   Data Management Planning Questionnaire for Data Managers and Information Technologists

This appendix, available at https://pubs.usgs.gov/publication/ofr20231069, includes a copy of the survey administered to data managers and information technologists. The survey consisted of multiple choice, multiple selection, Likert scale, and free-text questions.

Appendix 3.   Data Management Planning Questionnaire for Center Directors

This appendix, available at https://pubs.usgs.gov/publication/ofr20231069, includes a copy of the survey administered to Center Directors. The survey consisted of multiple choice, multiple selection, Likert scale, and free-text questions.

Appendix 4.   Data Management Planning Questionnaire for Program Coordinators and Bureau Approving Officials

This appendix, available at https://pubs.usgs.gov/publication/ofr20231069, includes a copy of the survey administered to Program Coordinators and Bureau Approving Officials. The survey consisted of multiple choice, multiple selection, Likert scale, and free-text questions.

Appendix 5.   Interview Questions for Researchers

This appendix, available at https://pubs.usgs.gov/publication/ofr20231069, includes a copy of the interview questions that were asked of selected researchers. Ten researchers were interviewed in total.

Appendix 6.   Interview Questions for Data Managers

This appendix, available at https://pubs.usgs.gov/publication/ofr20231069, includes a copy of the interview questions that were asked of selected data managers. Ten data managers were interviewed in total.

USGS