4 | Data Management Assessment and Planning Tools

ANDREW SALLANS AND SHERRY LAKE

INTRODUCTION: THE CHALLENGE OF DATA MANAGEMENT

Data is one of the hottest topics in recent years. In the academic world, we see continuous discussion of new initiatives for data-intensive research, of how institutions and disciplines should engage with "big data," and what new data skills are needed to remain competitive in a changing landscape. Much of this change is driven by advances in technology, leading to new opportunities for communicating, collaborating, and rethinking how research is done. Underneath it all, the fundamentals of managing research data become ever so much more important. Although the importance of managing research data is becoming better recognized in the academic environment, the tools and practices are still lagging and generally see slow adoption rates. That said, researchers and sponsors recognize that there is substantial risk due to reduced funding sources, missed opportunities from prior research projects, and the simple issue of insufficient knowledge and attention placed on research data management practices. This chapter will look at some approaches to addressing these issues, from our perspective as service providers in an academic research library.

In our experience, funding agencies provide general information about what the data management plan (DMP) should include, but they often are vague about requirements and provide few resources for researchers to consult when creating their DMPs. Without the proper training or background in data management and digital curation, researchers are apt to continue their current uninformed and incomplete data stewardship practices.

As data management service providers, we often see data management services as being driven by funder requirements, and more specifically, funder requirements for data sharing. As a result, the services that spark up around these requirements are rather limited in nature, often focusing foremost on the language and interests of the given funders, and less so on the idealistic or most pure aspects of data management best practices. This is understandable, given limited resources all around and limited experience, but in our view, this leads to subpar products later on down the line, and in the end, it does not really change behavior or improve practice all that much. In effect, in comparison to a teaching environment, it is much like "teaching to the test." Our data management service team has adopted the philosophy of focusing on best practices first and requirements second, hoping to have a longer-lasting impact upon behavior and skill development. As a result, our services and attitudes around tool development are guided by a long-term view, with a reality check to enable immediate application.

On the side of the researcher, there are similar barriers around improving research data management practices. Time, effort, and resources (both money and staffing) often are noted as reasons for poor management of data, but we also identify a lack of training, poor knowledge of best practices, and insufficient support or guidance as key factors. Most domains in academic research still lack formalized and standardized procedures for managing research data across the life cycle, as well as commitment to training new researchers in proper practices. It is unfortunately often a vicious cycle that keeps practices from improving.

In an effort to develop an understanding of how researchers at the University of Virginia (UVa) manage their research data, UVa Library's Scientific Data Consulting Group (UVa SciDaC Group) began a series of research data interviews. The goals of the data interview process included identifying common research data problems, identifying research support needs, and providing recommendations on improving data management. In practice, however, providing objective suggestions for data management practices proved to be troublesome. It was difficult to make reliable, customized recommendations and be objective in a timely fashion. In response to these challenges, the UVa SciDaC Group developed a system (DMVitals) to easily and objectively rate the current state of the researcher's data management practices.

Our experience with data management planning tools took a different path, and in this chapter, we will primarily discuss the philosophy surrounding the DMPTool (https://dmptool.org). First, however, we must touch on our tool's predecessor, the DMPOnline (https://dmponline.dcc.ac.uk/). DMPOnline was developed by the Digital Curation Centre (DCC) in the United Kingdom following recommendations in the *Dealing with Data* report (Lyon, 2007), an analysis of major UK funding requirements (Jones, 2010), and a period of initial experimentation with development of a "Checklist for a Data Management Plan" (Donnelly & Jones, 2011). As a web-based tool, the DMPOnline offers analysis of funding requirements, support in developing data management plans, and functionality for continued management of research data throughout the entire life of a project. Coverage of funders is mainly for UK agencies and some European Union organizations.

In response to the May 2010 announcement of new United States National Science Foundation (NSF) requirements for data management plans with all new proposals, we began collaborating with several other US research institutions on a US version of the DMPOnline. Although first envisioned as an expansion of the DMPOnline to include US funding agencies, it turned out that funding and research cultures in the United States and United Kingdom are different enough that it made more sense to develop a new tool while still maintaining collaboration between respective teams to share experiences and insights into best practices for data management planning, delivery of services, and anticipating needs in a rapidly evolving environment (Sallans & Donnelly, 2012). In the sections to follow, we offer insights into the development of DMVitals and the DMPTool, how they are employed, and the benefits they may bring to the delivery of research data management services.

DMVITALS TOOL

The DMVitals is a tool designed to take qualitative interview information and use it to systematically assess a researcher's data management practices in direct comparison to institutional and domain standards. Using the DMVitals, a consultant matches a list of evaluated data management practices with responses from an interview and ranks the researcher's current practices by their level of data management "sustainability." The tool then generates customized and actionable recommendations, which a consultant then

provides to the researcher as guidance to improve his or her data management practices. By design, the recommendations are far more objective, repeatable, and can be generated rapidly. The use of the DMVitals tool has helped our team expedite and standardize the data management consulting process.

Using best practice statements from UVa sources (Information Security, Policy, and Records Office and SciDaC Group guidelines) and the Australian National Data Service's (ANDS) long-term sustainability scoring model, the system compares the information collected during the data interview process with data management best practice statements. The model then further correlates the researcher's data management practices with the eight data management practice categories developed by the SciDaC Group: file formats and data types; organizing files; security/storage/backups; funding guidelines; copyright and privacy/confidentiality; data documentation and metadata; archiving and sharing; and citing data.

A key part of the tool is the data management (DM) sustainability ratios, which are created for each of the previously mentioned categories using the best practice statements. To provide a framework for defining and improving researchers' data management practices, the DM sustainability ratios are averaged to define a data management maturity level. These levels of maturity are a synthesis of the levels described by Crowston and Qin (2010) and the Australian National Data Service (ANDS, 2011), which are based on the Capability Maturity Model (CMM), typically used in software development (Paulk, Curtis, Chrissis, & Weber, 1993).

The data management maturity level is compared to the maturity levels as defined in the ANDS's *Research Data Management Framework: Capability Maturity Guide* (2011):

- Level 1: Initial (environment is not stable enough to support DM; few individuals have expertise; and infrastructure is disorganized)
- Level 2: Development (the researcher increasingly recognizes the lower level/easier best practices; DM process is under development)
- Level 3: Defined (the researcher is further defining his or her DM practices)
- Level 4: Managed (DM is seen as important at an organizational level and emphasizes coordination between the researcher and other organizational IT units)
- Level 5: Optimizing (the researchers are continually improving their data management practices; DM practices are not static)

The strength of the DMVitals tool is the creation of the DM report, which generates tasks customized to each researcher. These tasks can then easily be grouped into phases, creating a DM implementation plan for each researcher based on his or her personal data interview and subsequent information gathering. Combining this tool with assessment and planning methods helps to expedite the recommendation report process and provide valuable actionable feedback that the researcher can use immediately to improve the sustainability of his or her data.

UVa's Data Interview Initiative

The DMVitals tool evolved out of a need to systematically assess, score, and deliver mostly objective recommendations to a researcher following a start-up consultation interview. When we first began developing our DM services, we recognized that a data interview structure could help us develop a deeper understanding of how UVa's science and engineering researchers manage their research data while also initiating a discussion about how to simplify processes and improve practices. These interviews were constrained to 60 minutes and included the scientific data consultants, the subject librarian, and the researcher. With the data interviews we hoped to:

a. Identify common research data problems and needs,
b. Identify the types of digital data that are being created,
c. Identify communities and individuals who are under the most pressure from emerging grant regulations,
d. Identify potential partnerships for institutional repository data deposit, and
e. Develop opportunities to provide data management recommendations and training.

In creating our interview model we consulted the following models: the Data Audit Framework (DAF) for background reading (Jones, Ball, & Ekmekcioglu, 2008), the University of Oregon for consultation and information on implementation and buy-in (Westra, 2010), the University of Glasgow for interview questions and format (Ward, Freiman, Molloy, Jones, & Snow, 2010), and Purdue's Distributed Data Curation Center (D2C2) Data Curation Profile framework for question refinement (Carlson & Witt, 2007). Our interview protocol was based on the protocol from Wisconsin's Summary Report of the Research Data Management Study Group (Wolf et al., 2009).

Over the first two years of our services, we conducted about 25 data interviews. Through the process, we learned about research data practices at UVa, identified service needs and opportunities, and opened the door to consulting opportunities with researchers. Additionally, we encountered the dilemma of how to manage "unique" conditions of each research environment against common characteristics of DM within domains and institutional frameworks. In terms of support, we were having trouble customizing data recommendation requirements for each researcher.

We recognized a need to reduce the subjectivity and increase the speed at which we produced a report with recommendations. Additionally, we wanted to weigh all assessment factors from our interview, create actionable and repeatable recommendations, and address current DM conditions while showing paths for improvements. These needs are what led to the development of the DMVitals tool.

Development of the DMVitals Tool

The DMVitals tool is built using Microsoft Excel and consists of three types of worksheets: interview questions sheet, data management category sheets, and the data management report sheet. The interview questions sheet (Interview) contains the questions from the data interview. Each of the data management categories, as defined on the SciDaC Group website, are a sheet. In version 1.0 of DMVitals, we use five of the eight categories as sheets: file formats data types (FileFmtsDataTypes), organization of files (OrgFiles), security storage backups (SecStrgBackups), copyright privacy confidentiality (CopyrightPrivConfid), and data documentation metadata (DataDocMetadata). Each category sheet is populated with DM best practice statements for each category from UVa sources (Information Security, Policy, and Records Office (ISPRO), SciDaC Group guidelines) and the ANDS long-term sustainability scoring model. The category sheets also contain the calculated ratio of best practices statements. The third type of sheet, the data management report sheet (Report), is where the DM sustainability index ratios are displayed from the data management category sheets' ratio of best practice statements. This sheet also displays the data management maturity level (the average of the DM sustainability index scores) and the action statements for DM improvement (corresponding to best practices).

Interview Questions Sheet

The questions from the SciDaC Group's Data Interview Protocol are entered on the Interview sheet, one question per column. The current version maps questions from sections 2–5 (see Figure 1). Each question is then associated with one or more DM best practice statements. These DM best practices are listed under each question (or sub-question). Using the answers from the interview, each best practice was coded "yes," "no," or "null." "Yes" meant that the researcher was already doing that action; "no" meant that the researcher was not doing that action and "null" meant that best practice did not apply (for example, the best practice of "data is de-identified" in cases where data need not to be de-identified). "Null" is the default answer.

Figure 1. Interview sheet.

Data Management Category Sheets

Each best practice statement from the interview questions sheet is mapped to one of eight data management categories (file formats data types: FileFmtsDataTypes, organization of files: OrgFiles, security storage backups: SecStrgBackups, copyright privacy confidentiality: CopyrightPrivConfid, data documentation metadata: DataDocMetadata, funding guideline: FundingGuide, archiving and sharing data: ArchSharing, citing data: CitingData). Note that in this version, only five of the management categories are being used. The categories funding guidelines, archiving and sharing, and citing data are not used at this time. These will be added in a future version of DMVitals.

Each best practice statement from the interview questions sheet is categorized, given a sustainability level, and is then put in the corresponding column per one of the five sustainability levels (least sustainable, fair, satisfactory, good, and more sustainable). For ease of editing, each best practice statement is linked from its cell on the interview questions sheet to the sustainability level. The actual response to the best practice statement from the interview questions sheet—"yes," "no," or "null"—is also linked. The mapping and linking of each best practice is done for each of the categories on the corresponding category sheet.

To calculate the sustainability index ratio (current best practice to total-possible best practice), each sustainability level was given a "weight": least sustainable (1), fair (2), satisfactory (3), good (4), and more sustainable (5). The total number of "yes" responses (current practice) for each level is multiplied by the sustainability weight. These levels are then totaled and divided by the maximum number of best practices that apply, for that sustainability level, multiplied by the sustainability weight). The ratio for each category is then automatically recorded (via a link) on the data management report sheet as the sustainability ratio. See the screenshot in Figure 2 for the OrgFiles sheet for an example.

Data Management Report Sheet

The data management report sheet is comprised of three distinct sections: sustainability index, data management maturity level and action statements (see Figure 3). The top chart (sustainability index) shows the DM category and the resultant sustainability index (displayed as a percent—a ratio). The

	Least Sustainable (*1)		Fair (*2)		Satisfactory (*3)		Good (*4)		More Sustainable (*5)			
			NO	Using a consistent file structure					NO	Record every change to a file, no matter how small (Log files)		
			NO	Using file naming conventions			NO	Using file naming conventions for specific disciplines	NO	Uses File Version Control (perhaps SVN)	Total YES:	0
	NO	File changes are recorded					NO	Making original document "read only"	NO	Use Same Structure for Backups	Total Possible:	30
	NO	Versions are managed									Ratio	0
	NO	Files are organized and can be found										
	0	Total YES	0	Total YES	0	Total YES	0	Total YES	0	Total YES	0	Total
	3	Total Possible	4	Total Possible	0	Total Possible	8	Total Possible	15	Total Possible	30	Total

Interview | Report | FileFmtsDataTypes | **OrgFiles** | SecStrgBackups | CopyrightPrivConfid | DataDocMetadata

Figure 2. Data management category sheet—organization of files (OrgFiles).

actual ratios are linked from the corresponding data management category sheets. With five levels of sustainability, we divided the ratio values into five groupings—0–20 percent (Level 1), 21–40 percent (Level 2), 41–60 percent (Level3), 61–80 percent (Level 4), and >81 percent (Level 5)—and color-code the values using the colors on the data management maturity level scale (see the colors on Figure 3). This gives a visual view of how the researchers' current DM practices, per category, are ranked, according to the level of "sustainability." The ratios of the categories are averaged for a sustainability index. The average is also color-coded using the above percent groupings.

The bottom of the sheet contains a chart of action statements. The chart includes actionable recommendations targeting improvements for DM. Each best practice statement from the interview questions sheet has a corresponding action statement. The best practice a researcher is

not doing is marked with "X." These are the basis for DM improvement recommendations. The phases are customizable and can be moved around as the consultant sees fit. In our consulting approach, we avoid placing too many actions in any phase, which in turn may put researchers more at ease with improving practices.
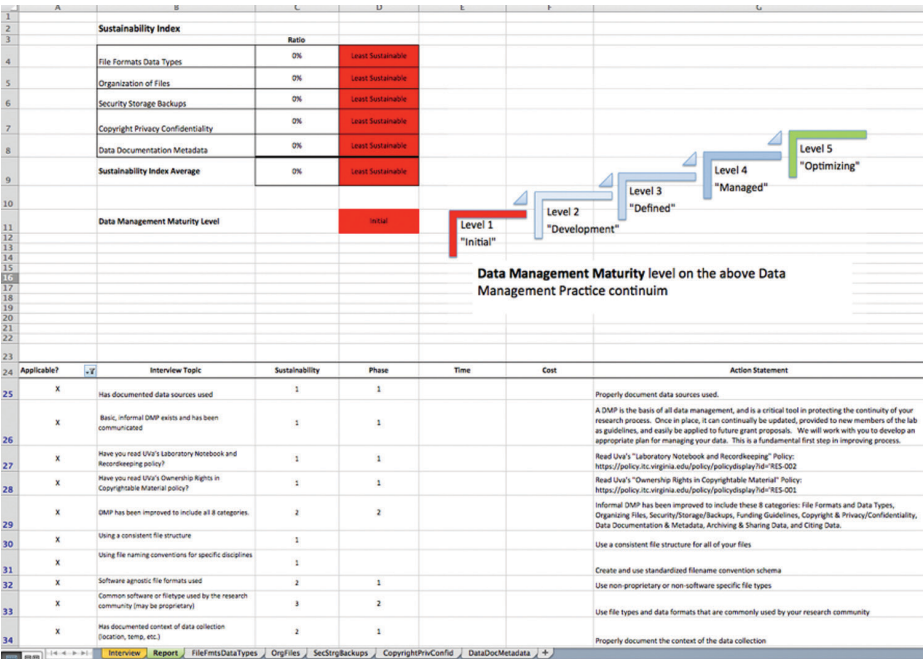
Figure 3. Report sheet.

Use of DMVitals Tool

Recommendations Report

The recommendations report is designed for distribution to researchers (see Figure 4 for an example). It begins with general information on DM and the goals of the report. The DMVitals report sheet provides the rest of the information that goes in the report. The sustainability index chart includes their data management maturity "grade." The chart is copied and pasted from the report sheet. The action statements are grouped into implementation phases (Phase 1: short-term; Phase 2: long-term; and Phase 3: future).
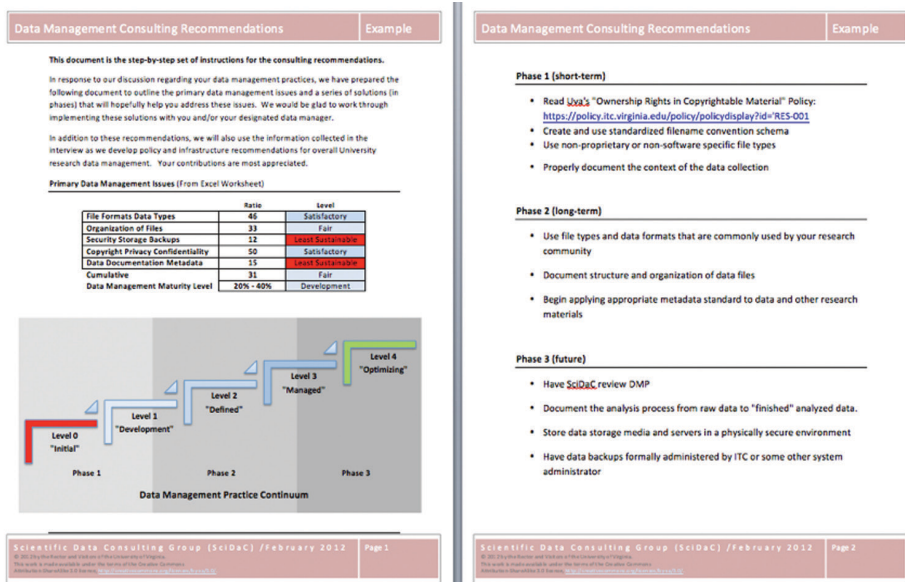
Figure 4. Data management consulting recommendations report.

Data Management Implementation

The next step in UVa's data interview workflow is to distribute the final report, with recommendations, and begin implementation with the researcher. We then sit down with the researcher to go over the recommendations and make adjustments on what actions are done in each phase (see Figure 5). The DMVitals tool is to be used throughout the implementation. As researchers improve their practices, their DMVitals score is recalculated and reflected in their new sustainability index.

The goal of repeating these steps—going back to the DMVitals, updating the best practices, and reevaluating the sustainability index—is for the researcher to obtain a data management maturity score of five "Optimizing," on the ANDS CMM for research DM. But as the model warns, level five is not the "final" level. Level five is labeled "Optimizing." At this level, researchers should be focusing on continually improving their DM practices.

Applying the DMVitals Tool at Your Institution

The DMVitals tool can easily be configured for your institution. It can be used with any data assessment tool. The main interview worksheet can be

customized using your assessment questions, local institution policies, and best practices. Ranking of sustainability can be adjusted per discipline or institution. The action statements definitely will require local customizations. These are the actions that your researchers need to do for your institution. Actions might include contacting specific service providers for support. Figure 6 shows the steps in the DMVitals workflow.

The DMVitals tool will continue to undergo improvement to further evaluate whether the scoring accurately represents improved practices. We recognize that the selection of assessment criteria, the calculation and weighting of scores, and how performance is communicated to the researcher must be done carefully in order to be effective as a communication tool with researchers. We plan for continued refinement of all steps, and we aim to produce a tool that offers value to both DM support providers and the researchers who receive support.
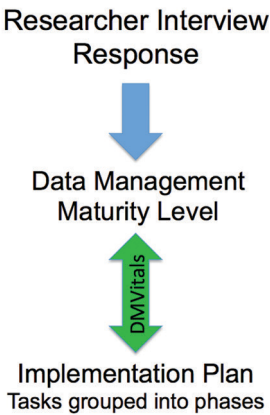
Figure 5. Data management implementation: DMVitals workflow.
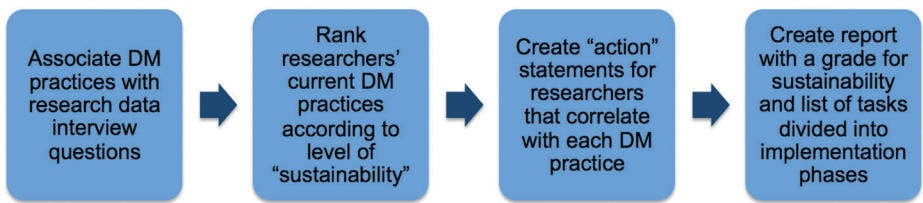
Figure 6. DMVitals modification flow.

At the date of writing this chapter, this is still a prototype tool, and we do not yet have metrics to assess the impact or benefit of its use upon effective delivery of services. We are in the early planning phases for an expanded version of the tool, and we likely will integrate it with other broadly used interview models.

DMPTOOL: GUIDANCE AND RESOURCES FOR YOUR DATA MANAGEMENT PLAN

Start Building a Framework

Development of the DMPTool began in early 2011, utilizing in-kind contributions of effort and resources from founding partners (see Acknowledgments), and led to the release of a production version in October 2011. During this first phase, software development was led by partners at the California Digital Library's UC Curation Center (UC3), content/workflow was led by partners at the University of Virginia Library, authentication/integration was led by partners at the University of Illinois at Urbana Champaign, and additional oversight and planning came from partners at Digital Curation Centre (United Kingdom), Smithsonian Institution, University of California-Los Angeles, and University of California-San Diego. Upon initial release, the tool included support for all specified requirements from the NSF, and it has since expanded to include requirements from the National Institutes of Health (NIH), the Institute of Museum and Library Services (IMLS), as well as other funders. The core philosophy behind the DMPTool approach is to: (1) help researchers create a quality and compliant DMP, and (2) inform researchers of resources and local support services available to assist with their data management planning needs.

Embrace a Community Support Model

Following successes in the first year, the team recognized a need to expand and embrace broader community participation. To facilitate the process, the DMPTool Alliance was formed in July 2012, and community building and outreach was placed as a high priority for all future efforts. In January 2013, the DMPTool partners were awarded a grant from the Alfred P. Sloan Foundation to fund further development of the DMPTool. In support of the grant, the DMPTool Alliance will be fully implemented over the course of 2013. This community model draws upon the expertise and resource contributions of original team members and institutions, but also provides structure for participation from the broader community of stakeholders. It includes decision-making processes, roles and responsibilities, terms of participation, and commitments to standards of operation within

the alliance, and is a central underpinning to the future strength of the DMPTool and its member community. At this point, the alliance plans to avoid a fee-based service model, and instead continue forward with an open-governance, shared-resource, and grant-funded approach.

From our perspective, the DMPTool offers a way for all information professionals to engage with researchers in the data management planning process. Recognizing that all organizations are at different points in the service development process, a support provider might use the service in the following ways:

1. An organization without staff allocated to data management support services can promote the DMPTool as the main service point, provide links to policies, storage, and other related institutional services and resources, and can begin to foster a discussion and effort at coordinating institutional support providers and services.
2. An organization with a small staff allocated to data management support services can gain some efficiency in repetitive or redundant tasks, allowing for more engagement with researchers.
3. An organization with a larger staff allocated to data management support services can use the DMPTool as a means of scaling and off-loading basic, repetitive tasks, allowing for more focus on high-value interactions, service architecture issues, and support for implementation of data management improvements.

In all cases, individual information professionals, teams, and overall organizations can benefit from approaching broad data management planning issues (i.e., understanding and tracking requirements, coordinating support across institutions, and sharing best practices) as a community, allowing for more time to focus on specialized local research support issues that cannot be initially addressed at scale. Technically speaking, institutions can take full advantage of DMPTool capabilities and community benefits by configuring their campus' single sign-on solution, typically Shibboleth. Institutions can then provide specific information for their researchers to customize their DMPs including help, suggested text, and additional resources.

Lastly, as a commitment to the community aspect of DMPTool, all software created by the project will continue to be made available under open-source licenses on a public code-hosting site, and DMPTool2 will include

development of an application programming interface (API) to allow the community opportunities to interface with other software. Following the conclusion of funded project work, ongoing management and governance of the DMPTool will remain community-oriented under the leadership of the DMPTool Alliance.

Approaches for DMPTool2

At the time of writing this chapter, we are nearly halfway through the development of DMPTool2. Through the community approach mentioned above, this new version places emphasis on meeting the needs of main user constituencies (i.e., researchers, librarians, funders, and institutions), with flexibility for expansion and integration of the tool in the future. The team hopes that the tool can serve as a point of convergence for sharing and dissemination of expertise and advice in good DM practices and solutions. For the information professional, this can be a support network for teams of one, and it can help to reduce the burden of individually rediscovering all of the best practices, tools, techniques, and resources that have already been tested elsewhere. For institutions, this approach may offer an enterprise view of successful and failing DM practices, allowing for opportunity to improve and plan resources more effectively.

DMPTool2 is expected to include the functions detailed in Table 1. Although these are the projected functions, they will be revisited throughout the duration of the project via outreach, focus group testing, and evaluation of intermediate prototypes, and are thus subject to change.

Metrics and Impact

Although supporting researchers in data management planning is the core function of the DMPTool, we also recognize that attention to evaluation of the service and ongoing assessment of impact is of equal importance. This is true for operational considerations and use of resources within the information organization, but more importantly as a means of showcasing value to those whom we serve. There are many ways to approach metrics and impact in this case, and we will cover a few here.

Foremost, there are "low-hanging fruit" opportunities. Through Shibboleth authentication and basic database queries of usage, the team has been able to easily track basic web use statistics since the beginning.

New Function	Benefit
Granular modeling of plan templates	Plan templates can be refined with a hierarchical structure, which is important as plan requirements become more comprehensive and complex.
Granular modeling of institutions	Institutions can be represented in any or all of three roles: as a funder, as a researcher's affiliation, and/or as an institution with its own DMP requirements.
Role-based user authorization	Users can hold various roles: plan creators, collaborators, institutional/organizational administrators, and tool administrators, with differing authorizations and capabilities.
DMP life cycle management	Users can track the changes as a plan passes through the major stages of its life cycle, including creation, editing, submission, evaluation and approval or rejection, revision and publication.
Organizational planning activities	Institutions/organizations can define and implement planning activities, with consideration of local resources and needs.
Enhanced search and browse	Users can both perform keyword searches and browse publicly available plans, providing for enhanced discovery.
Institutional branding	Two aspects: (1) hosting neutrality of the DMPTool, that is, removing association of the tool with the California Digital Library via the current domain name dmp.cdlib.org; and (2) increased online institutional branding within the tool's user interface.
Search and reporting for business intelligence	Institutional administrators can mine data on plans, with the ability to filter plans based on their state, users, institutional role and other relevant plan properties.
Advanced administrative interface	Administrators can easily add new funders, requirements, and institutional resources via the enhanced configuration interface.
Collaborative plan creation	Multiple users across institutions can access and/or edit the same plan.
Open API	The tool will be more easily integrated into local automated workflows.

Table 1. Overview of DMPTool2 functions and benefits. Table borrowed from grant proposal and Alfred P. Sloan Foundation grant project documentation.

The team has tracked usage by number of users, number of plans created, and number of institutions represented. During the first 18 months since launching the DMPTool, we saw over 5,000 unique users, over 4,500 plans created, and representations from over 700 institutions. There has been

continuous, constant growth of these statistics over the entire course of the service. These use statistics, along with several awards and grants (see Notes), have been helpful for raising awareness amongst particular communities, and for attracting additional resources for growth and continued community building to create a center of mass. While we find these trends encouraging, they do not, however, say much about the qualitative aspects of the service or the impact that data management planning has upon success in research, compliance with requirements, or more openness of research data.

The DMPTool2 project aims to improve what we can learn from use of the service. Current plans aim to look at factors such as adoption of new functionality by existing users, recruitment of new users, and recruitment of new developers contributing to the code base. These elements hopefully will shed light on the best places to focus outreach and educational efforts. Likewise, the project will dig deeper to identify which characteristics may influence higher institutional use rates, and consequently (if possible), better DM practices amongst researchers.

One area that currently looks promising is the institutional adoption demonstrated by integration of DMPTool with local Shibboleth authentication. As one might suspect, taking the step of setting up authentication often also leads to broader promotion at the institutional level, which leads to more opportunities to pull together institutional DM support providers into one service point, which leads to better awareness and informed decision-making by researchers, and lastly, which will hopefully result in better data management plans and practices. DMPTool usage data amongst Association of Research Libraries (ARL) institutions indicates this type of trend, but more thorough analysis will need to be done in future project development before releasing any conclusive results.

Although the factors outlined above are important in evaluating value and impact, we also recognize that there will need to be continued attention to how services and tools like these impact broader issues in data management planning. Some questions might include:

- Does data management planning support provided by the DMPTool lead to improved data management practices and outcomes?
- Does use of the DMPTool correlate with funding proposal acceptance?
- Does the DMPTool actually save the researcher time and effort during proposal writing?

- Does the DMPTool improve the efficiency or effectiveness of institutional data curation service providers?
- Does use of DMPs lead to increased data publication, citation, and sharing, and if so, does that sharing enable avenues of scholarly research and discourse that might not otherwise occur?

Some of these questions may be possible to answer through the planned addition of institutionally oriented business intelligence and mining functions in DMPTool2. For information professionals, this may mean gaining a better understanding of research practices and behaviors, anticipating demands on repositories, or gauging how well researchers understand intellectual property ownership and responsibilities, all gained with little to no interruption of the researcher. Having access to such information may allow for more intelligence, user-oriented design of services, infrastructure, policies, and better anticipation of staffing needs, with minimal cost, effort, and possibly more candid responses. Although these functional needs have been expressed by a number of different institutions, and the team recognizes their value, there are still a variety of policy concerns that will need exploration.

While it may not yet be possible to demonstrate true correlation between the use of the DMPTool and improved data management practices, we believe that use of the DMPTool is a straightforward and easy way for libraries to become more active and engaged in the DM conversation. Our hope is that the structure of the tool facilitates a positive DMP experience for researchers, and leads to learning and insights by librarians and other support providers. It is very easy to get started. Simply visit http://dmptool.org to find more details on setting up an institutional profile to support your researchers.

CONCLUSION

Given the points raised here, we argue that research data management assessment and planning tools can add tremendous value to the services provided by information professionals to the research community today, and can become a true game changer. With this opportunity come expectations for very high-quality products and services, which are tailored to the needs and culture of the academic research community.

Lastly, we believe that it is important to stress that improvement of data management practices must begin somewhere. Going through the pro-

cess of documenting existing practices affords everyone an opportunity to reflect and determine the best path toward improvement. Along these lines, having an active, operationally focused DMP will enable far easier and better planning for funding proposal data management plans, whether broad in scope or limited to data sharing interests. We hope that additional information professionals can employ the DMVitals and DMPTool to enhance initial data management support services, providing standardization, consistency, and scale.

ACKNOWLEDGMENTS

NOTE

The DMPTool has received a number of awards and grants during the period of development, including (chronologically):

- Library of Congress' Top 10 Digital Preservation Developments of 2011 (January 2012, http://blogs.loc.gov/digitalpreservation/2012/01/top-10-digital-preservation-developments-of-2011/)
- Larry L. Sautter Golden Award for Innovation in Information Technology (January 2012, http://www.cdlib.org/cdlinfo/2012/07/03/dmptool-wins-sautter-award/)
- $149,070 grant from the IMLS National Leadership program (September 2012, http://www.imls.gov/news/national_leadership_grant_announcement_2012.aspx)
- Finalist for Digital Preservation Coalition's Digital Preservation Award for Research and Innovation 2012 (December 2012, http://www.dpconline.

org/advocacy/awards/2012-digital-preservation-awards/928-finalists-
2012-research-and-innovation)

- Grant for $590,000 from the Alfred P. Sloan Foundation (January 2013, http://www.cdlib.org/cdlinfo/2013/01/14/cdl-and-partners-receive-sloan-funding-to-enhance-dmptool-features-reach-out-to-community/)

REFERENCES

Australian National Data Service (ANDS). (2011). *Research data management framework: Capability maturity guide*. Retrieved from http://ands.org.au/guides/dmframework/dmf-capability-maturity-guide.html

Carlson, J., & Witt, M. (2007). *Conducting a data interview*. Presented at the 3rd International Digital Curation Conference, Washington DC. Retrieved from http://docs.lib.purdue.edu/lib_research/81

Crowston, K., & Qin, J. (2010). A capability maturity model for scientific data management. *American Society for Information Science and Technology Annual Meeting*. Pittsburg, PA. Working Paper available: http://crowston.syr.edu/content/capability-maturity-model-scientific-data-management-0

Donnelly, M., & Jones, S. (2011). Checklist for a data management plan. *Digital Curation Centre*. Retrieved from http://www.dcc.ac.uk/resources/data-management-plans/checklist

Jones, S. (2010). Summary of UK research funders' expectations for the content of data management and sharing plans. Retrieved from http://hdl.handle.net/1842/3374

Jones, S., Ball, A., & Ekmekcioglu, C. (2008). The data audit framework: A first step in the data management challenge. *International Journal of Digital Curation*, *3*(2), 112–120. http://dx.doi.org/10.2218/ijdc.v3i2.62

Lyon, L. (2007). *Dealing with data: Roles, rights, responsibilities and relationships*. Consultancy Report for the JISC. Retrieved from http://www.jisc.ac.uk/media/documents/programmes/digitalrepositories/dealing_with_data_report-final.pdf

Paulk, M. C., Curtis, B., Chrissis, M. B., & Weber, C. V. (1993). *Capability maturity model for software, version 1.1* (CMU/SEI-93-TR-024). Pittsburgh, PA: Software Engineering Institute, Carnegie Mellon University. Retrieved from http://www.sei.cmu.edu/reports/93tr024.pdf

Sallans, A., & Donnelly, M. (2012). DMPOnline and DMPTool: Different strategies towards a shared goal. *International Journal of Digital Curation*, *7*(2), 123–129. http://dx.doi.org/10.2218/ijdc.v7i2.235

Ward, C., Freiman, L., Molloy, L., Jones, S., & Snow, K. (2010). *Incremental scoping study and implementation plan.* Project Report. University of Cambridge, Online. Retrieved from http://eprints.gla.ac.uk/54623/

Westra, B. (2010). Data services for the sciences: A needs assessment. *Ariadne, 64.* Retrieved from http://www.ariadne.ac.uk/issue64/westra/

Wolf, A., Simpson, M., Salo, D., Flee, D., Cheetham, J., & Barton, B. (2009). *Summary report of the Research Data Management Study Group.* Madison: University of Wisconsin. Retrieved from http://digital.library.wisc.edu/1793/34859