

5-15-2013

# Building an Online Data Management Plan Tool

Michele Reilly

Central Washington University, reillym@cwu.edu

Anita R. Dryden

University of Houston - Main

Follow this and additional works at: <http://digitalcommons.cwu.edu/libraryfac>



Part of the [Higher Education Commons](#), and the [Scholarly Communication Commons](#)

---

## Recommended Citation

Reilly, Michele and Dryden, Anita R., "Building an Online Data Management Plan Tool" (2013). *Faculty Scholarship from the Library*. Paper 1.

<http://digitalcommons.cwu.edu/libraryfac/1>

This Article is brought to you for free and open access by the Library at ScholarWorks@CWU. It has been accepted for inclusion in Faculty Scholarship from the Library by an authorized administrator of ScholarWorks@CWU.

## Building an Online Data Management Plan Tool

Michele Reilly, Anita R. Dryden

---

Reilly, M, Dryden, AR. (2013). Building an Online Data Management Plan Tool. *Journal of Librarianship and Scholarly Communication* 1(3):eP1066. <http://dx.doi.org/10.7710/2162-3309.1066>

© 2013 by the author(s). This open access article is distributed under a Creative Commons Attribution License, which allows unrestricted use, distribution, and reproduction in any medium, providing the original author and source are credited.

JLSC is a quarterly journal sponsored and published by Pacific University Library | ISSN 2162-3309 | <http://jlsc-pub.org>

# Building an Online Data Management Plan Tool

Michele Reilly *Head of Digital Services, University of Houston*

Anita R. Dryden *Digital & Web Projects Fellow, University of Houston*

## Abstract

Following the 2011 announcement by the National Science Foundation (NSF) that it would begin requiring Data Management Plans with every funding application, the University of Houston Libraries explored ways to support our campus researchers in meeting this requirement. A small team of librarians built an online tool using a Drupal module. The tool includes informational content, an interactive questionnaire, and an extensive FAQ to meet diverse researcher needs. This easily accessible and locally maintained tool allows us to provide a high level of personalized service to our researchers.

## INTRODUCTION

On January 18, 2011 the National Science Foundation (NSF) began requiring supplemental data management plans (DMP) as part of every funding application. This new requirement created a great deal of confusion and anxiety in the research community, including among researchers at the University of Houston. At the University of Houston Libraries, the first call for help came on January 15th of that year from a faculty researcher requesting assistance with her DMP. She had been uncertain as to what to include in her DMP and was seeking guidance from other sources. She solicited help from her colleagues but wasn't confident with the results, which consisted of a three sentence example that a fellow researcher had given her. We (the researcher and librarians) didn't feel that this was adequate and weren't sure that it would satisfy the NSF. After many phone calls, emails, and further online research we developed a DMP that we thought would better fulfill the requirement.

For this researcher and others, the new DMP requirement

was initially irritating and confusing because it requires researchers to think differently about the data they create. It compels researchers to think beyond what their data means to them and consider what it might mean to someone else. It introduces the concepts of the fragility of data, the interoperability of data, and the sharing of data. Conveying these concepts to our faculty researcher helped her understand the need and benefit of a DMP—not only to meet the NSF requirements, but to guide her in effectively managing her data for future use.

That first faculty member request spurred the University of Houston Libraries' Head of Digital Services to explore simple ways to help our researchers develop a DMP through the most expedient and user-friendly approach possible. Our guiding question was one that is shared by many libraries: What can the library provide that will help researchers develop data management plans when they don't even know what those plans should contain? This paper describes how the University of Houston Libraries answered that question through the creation of a concise, interactive online tool.

Received: 01/22/2013 Accepted: 04/29/2013



© 2013 Reilly & Dryden. This open access article is distributed under a Creative Commons Attribution 3.0 Unported License, which allows unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## LITERATURE REVIEW

Although many researchers (and librarians) were surprised by the new data management requirements, libraries have been preparing for the new landscape of research data curation for quite some time. Much of the early literature comes from the United Kingdom, where government mandates for research data sharing gained traction a few years earlier than in the United States, but a number of individuals and institutions in the U.S. were actively pondering librarian roles in data management early in the millennium. For example, in 2006 the Association of Research Libraries issued a call for the creation of long-term data stewardship frameworks (ARL 2006).

Beyond organizations like ARL, libraries that were early leaders in exploring data management roles included Purdue University (Brandt 2007, Witt 2008, Witt et. al 2009, Cragin et. al. 2010), Johns Hopkins University (Choudhury 2008, 2010), and Cornell University (Steinhart 2007, Steinhart et. al. 2012). More recently, a number of American universities have undertaken local studies of the needs of researchers on their campuses in order to better understand opportunities for library involvement in data management. These case studies, including one at the University of Houston (Peters & Dryden, 2011), have shown a wide range of attitudes and levels of preparedness among faculty researchers (Bracke 2011, Latham & Poe 2012, Parham et. al. 2012). Most of these case studies involve surveying or interviewing faculty members at the institution, but Lage et. al. (2011) offer a novel approach towards the creation of “personas” for researchers with different needs and attitudes related to data management.

As evident from the early leaders (e.g. Purdue), much of the conversation around research data management has come from large research institutions, but other types of universities are also beginning to research their institutional needs. Scaramozzino, et. al. (2012) has identified strong roles for librarian involvement in research data curation even at traditionally teaching-centered universities, particularly in terms of educating researchers about data management practices and requirements. Similarly, Shorish (2012) provides suggestions for small steps that librarians at masters and baccalaureate granting institutions can take to engage their faculty around data management issues. Even some very small liberal arts colleges have begun to explore data curation

opportunities, notably Mount Holyoke and the other members of the “five colleges” consortium (Goldstein & Oelker, 2011).

As a necessary adjunct to exploring the role of libraries, attention has also begun to be paid to *librarian* preparation for data management roles. Harris-Pierce and Liu (2012) question the adequacy of North American LIS degree programs for preparing librarians to enter these emerging roles, and find that while many new courses have been created to respond to this demand, there are still far too few. Cox et. al. (2012) discusses potential curricula for training existing science librarians in competencies needed to support research data curation and management. They suggest an 8-part curriculum based on previously identified roles for librarians involved in data management. Steinhart and Qin (2012) discuss their mentoring collaboration with a newly developed certificate program in eScience Librarianship at the University of Syracuse.

Of course, the preparation of libraries and librarians for data management roles only addresses one side of the issue; studies such as Johnston and Jeffryes (2013) have also explored the need to prepare students in STEM disciplines to meet new data requirements and expectations in their current and future research. This is especially important because, although the concepts of public access to federally funded research data and data management plans generally have been around for some time, researchers seem to have been largely unprepared to meet the 2011 NSF mandate and other similar requirements. In fact, according to a Cornell University Data management survey; “53% [*of respondents*] would be interested in any sort of guidance, including consultation,” for writing a data management plan in support of an NSF grant application (Steinhart, et. al., 2012). In the same study, Cornell also found that, “considerable confusion exists as to what ‘counts’ as data, even among researchers who are likely among their discipline’s experts.” Much of this confusion may stem from a lack of clarity about funding agency requirements for data management planning. Another Cornell study conducted in the summer of 2012 (Dietrich, et. al.) found that no single agency’s policy addressed all elements of what should be included in a Data Management Plan. Four of the evaluated funders appeared to have no policy at all; though the study did find some common requirements that pertained to general data management activities.

Cornell's findings illustrate the excellent opportunity that librarians have to educate researchers about the organization of their data and the role that data management planning plays in supporting long-term access to their data, the possibility of reuse of their data, and the potential for increase in citations to their research. These data management benefits can lead to better science, enhanced returns from publicly funded research, and improved linking between datasets. The opportunity both to provide a valuable—and needed—service to researchers and also to improve the overall accessibility and impact of research data is one that librarians should continue to explore and embrace.

## INSTITUTIONAL CONTEXT

As illustrated earlier, at the University of Houston Libraries we have found that creating a DMP is daunting for some researchers. Junior faculty and experienced researchers alike are often confused by what data management is, why they need to plan for it, and what should be included in a plan. Exacerbating this confusion, and making DMP assistance even more vital, is the fact that researchers have

been seeking grant opportunities far more vigorously than they have in the past due to the University of Houston's drive to become a Tier 1 research institution (University of Houston, 2010).

As grant activity expanded, the number of phone calls and emails we received quickly made it clear that researchers needed assistance in developing their DMPs. Many researchers were leaving their DMPs to the last minute, which made it critically important to be able to provide guidance that was understandable, easily navigable, and thorough. Initially, the Head of Digital Services considered designing a template, but discovered that there were templates available online already. Based on the calls received from researchers, it was apparent that they had questions that no available template was able to answer. Furthermore, templates did not seem able to provide the desired level of interactive assistance. An interactive web form that could help generate DMPs was identified as the best solution to the problem.

We were aware that there were other, more comprehensive, tools in development (see *Tool Review*), but none were

### TOOL REVIEW



**Figure 1. DMP Tool** (<https://dmp.cdlib.org/>)

Several universities and organizations came together to produce the DMPTool (California Digital Library 2013, Figure 1). These include: The University of California Curation Center (UC3) at the California Digital Library, DataONE, Digital Curation Centre (UK), Smithsonian Institution, University of California, Los Angeles Library, University of California, San Diego Libraries, University of Illinois, Urbana-Champaign, and University of Virginia Library.

Launched in October 2011, the DMPTool is designed to help researchers meet the new funder requirements by allowing them to create funding agency specific data management plans. It offers a step-by-step guide and instructions for generating a DMP. Features include:

1. Account creation. By creating an account the researcher can save a plan and come back to it later. Very useful if the researcher has unanswered questions at first log in.
2. Funding agency specificity. When the researcher creates a new plan, he/she chooses the appropriate foundation/agency/directorate for the grant.

The DMPTool takes the researcher through a 5 step process for creating a plan. Different information is required at each step: Roles and responsibilities, expected data, period of data retention, data format and dissemination, and data storage and preservation of access. After filling out this information, the researcher is given the opportunity to export the DMP to plain text or rich text (Microsoft Word). The format is clean and in a manner that the NSF prefers. The researcher can also copy and paste the information from the web page. (*cont'd*)

This tool is very user friendly, especially if the researcher knows what metadata is, where data will be stored, and how data will be shared. The funder requirement link pulls up the link to the funder and a funder requirements template and lets the researcher know whether the DMPTool supports that particular funder. The help section offers a step-by-step guide to using the tool and there is also an interesting Demo video.

Including information about local resources and services for researchers could be problematic. The only way to modify the form to make it institution specific is to have the institution become a contributing member. Some institutions may not be willing or may not have the resources to manage or maintain an institution specific instance.



**Figure 2. DMP Online** (<https://dmponline.dcc.ac.uk/>)

DMP Online was produced by the UK's Digital Curation Centre (2013, Figure 2). This tool helps research teams develop European funding agency required DMPs, but it works for U.S. funders also. The researcher is asked a series of questions; for example, regarding the stage in the application process the research is currently in and the type of funder to which the researcher has applied. If U.S. Funders is selected, the tool will ask the researcher to select the relevant funder and the only option is NSF-generic. The questions after this are all based on the researcher's answers, which is helpful since DMP Online then ensures that funder specific requirements are addressed.

For example, if a researcher is applying to one of the funders that make specific data-related conditions at the application stage, he/she will see the funder's requirements in the left-hand column. These have been mapped to the appropriate clauses in the DCC Checklist for a Data Management Plan (Digital Curation Centre, 2011). By answering the questions presented, the researcher should meet the funder's requirements.

At the present time any funders that do not have generic data-related conditions at the application stage are not currently mapped. The DCC hopes to work with all of the major funders to create accurate mappings that the funders can approve. Once the researcher indicates where his/her funding is coming from, a much longer series of questions is presented that covers all aspects of data management. These questions are, again, taken from the Checklist for a Data Management Plan (Digital Curation Centre 2011). The researcher has the option of answering as many or as few of the questions as desired within the section and can add or remove questions. For many of the questions, additional information is available by clicking the information button (i) on the right-hand side of the page.

The completed DMP can be exported in PDF, HTML or CSV formats. It is at this stage in the process that the researcher can reorder sections and can include or exclude sections. The DMP Online tool is very versatile and has some helpful features. At this time the only "US funders" template is for the National Science Foundation (generic). One drawback is that there isn't additional explanatory information for each question and the authors could find no example text. Additionally, this tool does not allow the inclusion of information about local resources and services that could be useful for researchers.



**Figure 3. IEDA Data Management Plan Tool** (<http://www.iedadata.org/compliance/plan>)

Launched in January 2011, the IEDA Data Management Plan Tool (Integrated Earth Data Applications 2013a, see Figure 3) is a simple one-page web form. The researcher answers the questions, clicks on 'create PDF', verifies the information, and the process is done (producing a tidy PDF). There was no help, explanations, or definitions, only links to the NSF requirements and a helpful suggested repositories list.

In April 2012, Version 2 was launched with enhanced features: login, re-use of previously made DMPs, associating DMPs with funded NSF awards, an IEDA DMP Tool FAQ, and a preview of the Tool. The login is tied to the GeoPass (Integrated Earth Data Applications 2013b) system, which users must register for if they are not existing members of the organization.

widely available at the time, and there was no indication of when these tools would become available. Because we wanted to offer better guidance to University of Houston researchers before the next grant cycle began, we decided that waiting for these tools would be unproductive, and it would be more efficient to develop an in-house solution.

It was determined that our online form would be researched, developed, and executed by a very small group. Using a small working group made it easier to be more agile and responsive to changes and technical issues, which ultimately helped the project go swiftly and smoothly. The working group was composed of the Head of Digital Services and a librarian fellow from the Libraries' Systems department; the systems librarian provided the technical expertise required to build the web form while the Head of Digital Services provided the content. Within three

months of the NSF announcement the team had designed a web form that would email researchers a draft DMP based on their answers to a questionnaire. This draft plan is in a format that is concise, easily readable, and includes what we determined the granting agencies wanted to see in a data management plan.

## FORM OVERVIEW

The University of Houston Libraries DMP form (<http://info.lib.uh.edu/p/dmpform>) consists of three separate web pages: an introduction/instruction page, the questionnaire, and an extensive FAQ. The first page of the form is informational and explains the purpose and the elements of a DMP and provides links to further information on this topic (Figure 4). It goes on to discuss how the form is assembled, what researchers can expect to

Figure 4. Data Management Information Page

UNIVERSITY of **HOUSTON** | LIBRARIES

M.D. Anderson Library  
Today's Hours 7:00am - 9:45pm

Home > Services >

## Data Management

### What is a Data Management Plan?

National Science Foundation (NSF) proposals submitted or due after January 18, 2011, must include a data management plan as a supplementary document. This data management plan cannot be more than two (2) pages in length and be labeled "Data Management Plan". The plan describes how the researcher will disseminate, share and store data, reports and information acquired within the grant period. For more information and requirements about this initiative and directorate initiatives see: <http://www.nsf.gov/bfa/dias/policy/dmp.jsp>

### Need help writing your data management plan?

This researcher populated template serves as an aid in fulfilling the data management plan (DMP) requirement mandated by funding bodies. Examples, guidance and suggestions for each question are included in the [FAQ](#) to assist you in decision making and DMP writing. The result will be a working, customizable data management plan emailed to you.

To start creating your data management plan, answer as many of the questions on the following pages as apply to your circumstances. We strongly suggest that you [print the questions](#) before filling out the form. This will allow you the opportunity to formulate your answers, look up directorate requirements, and familiarize yourself with unfamiliar terms. You will not be able to save your work in progress while completing the form, so please ensure that you have your answers at hand.

Remember your DMP can be no more than 2 pages so please keep your answers concise. If you stick to answers of 1 or 2 complete sentences per question, the DMP that is produced by this form should be no more than 2 pages, after you take out the email addresses and the introduction at the top.

Upon clicking submit, an email will be sent to the address you provide that contains your answers with the question text removed. Cut and paste the text starting with **Data Management Plan** into a text document to finalize your plan. You will want to modify your answers to form a narrative and to reduce it to 2 pages if necessary. Providing answers in narrative form originally should help in keeping the rewriting to a minimum. Your completed DMP will have section headers followed by narrative paragraphs. We suggest you leave the numbered header at the beginning of each section. [See an example](#) of a numerically formatted DMP.

If you need further assistance please contact Michele Reilly (713-743-8995) at the UH Libraries.

[Data Management Plan Form](#)

Permalink: <http://info.lib.uh.edu/p/data-management-plan>

### Data Management Resources & References

- [National Science Foundation -- Data Management Plan Requirements](#)
- [California Digital Library \(CDL\)](#)
- [Data Curation Centre \(DCC\) DMP online DMP tool](#)
- [Inter-University Consortium for Political and Social Research \(ICPSR\), University of Michigan](#)
- [MIT Libraries](#)
- [Research Cyberinfrastructure, UC San Diego](#)

© 2012 The University of Houston, 114 University Libraries, Houston, TX 77204-2000 713-743-1050  
[Website Feedback](#) | [Policies](#) | [Giving to the Libraries](#) | [Contact Us](#) | [Maps & Directions](#) | [Site Map](#) | [Mobile Site](#)  
[UH Home](#) | [UH System](#) | [State of Texas](#) | [Emergency Site](#)

generate as a result of completing the form, and the best way to finalize their plans. This introduction/instruction page was essential in providing guidance and context to the researcher on how to best answer the questions to ensure the least amount of editing in the end.

From this information/instruction page the researchers are able to preview and print the questions. This allows them to formulate their answers, look up terminology they may not be familiar with, and ask thoughtful questions about designing their DMP. Printing the questions in advance can guide the researchers in their writing and also allows them to create their DMPs without using the web form if they choose.

The questionnaire asks a series of questions and provides space for the researchers to fill in their answers (Figure 5).

It is organized into five sections of thematically similar questions. Each section begins with introductory text about the questions in that section as well as links to the FAQ and contact information for the Head of Digital Services. The sections are:

- Types of Data
- Data and Metadata Standards
- Policies for access/sharing and provisions for appropriate protection/privacy
- Policies and provisions for reuse and redistribution
- Plans for archiving and preservation of access

Upon submitting the form, the researcher receives an email containing the answers to the questionnaire with

**Figure 5. Data Management Plan Form: Example Page**

Home > Services >

## Data Management Plan Form

**V. Plans for archiving and preservation of access**

Arrangements must be made for long-term storage and preservation of your data. DMPs must include plans for archiving data, samples, and other research products and preservation of access to them. [Why is this important?](#)

For more information and examples, see the [FAQ](#).

**5.1 There are several places that this data can be stored. Please choose the one that best fits your needs:**

Personal responsibility

UH Libraries

Professional Society or research entity or data center

Other

No plans to preserve or provide access to materials

If you have chosen the library as where you will keep your data please copy and paste the examples from the [FAQ](#) as your answer to the following questions. There is a cost attached to this option with the exception of final reports and their accompanying addenda. Find out what the [NIH and NSF](#) have to say about funding your data management.

**If you have chosen an archive, repository, data center or central database in which to deposit your data please indicate which one and who will be responsible for maintaining data within this system:**

**5.2 Please explain your strategy for [maintaining, curating and archiving](#) your data?:**

**5.3 Please explain the preservation plan and back up strategy of your long-term data:**

**5.4 Along with your data there will be accompanying documentation. (i.e.metadata, references, reports, research papers, fonts, bid proposals, reports, etc.) Please list any of these that you think you will be generating:**



the question text redacted. If the researcher has filled out the form in short narrative sentences, it is only necessary to then copy and paste the body of the email into a Word document. Any needed edits can be made at that time. The average email response is approximately one and a half pages in length and after any edits the resulting DMP comes in under the two page limit that many agencies request.

We created a companion to the form which we called the DMP FAQ (Figure 6). This was a separate web page which provided extensive detail for each question on the form, including sample answers and definition of terms that might be unknown to researchers. An extensive amount of research was put into the FAQ to provide researchers with the most current information possible. This FAQ is frequently audited and updated to ensure accuracy.

An increasing variety of funding sources are requesting DMPs from researchers. For this reason, we chose to make our form granting agency and directorate agnostic. We felt that offering a general form would help the greatest number of researchers. On those questions where directorates have made specific requirements, the researcher is directed to the FAQ.

## FORM TECHNICAL NOTES

Our DMP web form was built in-house on our Drupal-based library website. We explored other options for creating the form, including third-party hosted solutions, but decided that building the form within the main library website would provide us with the greatest ability to customize the form itself. It also provides a better user experience through the more consistent look and feel of

**Figure 6. Data Management Plan FAQ: Example Page**

UNIVERSITY of **HOUSTON** | LIBRARIES

M.D. Anderson Library  
Today's Hours 7:00am - 9:45pm

Home > Services > Data Management Plan Form >

## DMP FAQs/Definitions

### 1.1 Description of data

Providing a brief overview of the data to be collected helps reviewers to understand the scope, nature and scale of the data that will be generated.

**Example: This project will generate data describing political party organizational development in 15 countries.**

What types of data are included?

All research data that were formally defined as "the recorded factual material commonly accepted in the scientific community as necessary to validate research findings" by the U.S. Office of management and budget (1999).

**Example: Examples of data include party membership size, gender balance of legislative delegations, national party income and expenditure patterns, and rules for selecting legislative candidates and party leaders.**

### 1.2 Data Capture or Creation

Describe how you plan to create this data or capture it using software or equipment.

**Example: Data will be uploaded into a spreadsheet from field notes and seismic measurement equipment.**

### 1.3 Existing Data

Only answer this question if you will be using existing data.

**Example: The data for this project will be culled from existing data but is difficult to obtain by standard channels.**

## 2 What is a metadata standard?

A simple metadata schema contains fields for TITLE, CREATOR, DESCRIPTION, DATE, RESPOSITORY, SOURCE, USE & REPRODUCTION. It is highly advised to include more information than this in your metadata schema. See industry standards below.

Examples of existing metadata standards:

- [Dublin Core Metadata Initiative – Metadata Basics.](#)
- [Metadata Research Center \(MRC\)](#)
- [Biodiversity Information Standards TDWG\(a list of several metadata standards for science research data\)](#)

What metadata standard does the UH Libraries use?The library uses Dublin Core. See above for link to the Dublin Core website. It is perfectly acceptable to create your own metadata schema if you do not feel that any existing schema meets your needs.

### 2.1 What is a file format?

A file format is the extension after the file name, e.g. filename.jpg, filename.txt, filename.xlsx.

**Example: The file formats that this project will generate are as follows: .txt, .docx, .xls, etc.**

our main website.

The form was built using the “Webform” module of the Drupal content management system (Figure 7). This is a fairly straightforward tool for building web forms with an easy-to-understand user interface. The Webform module allowed us to break up each question into three different components: question text, “helper” text to further clarify the question, and example text.

We used these text components in order to keep the questions as straightforward as possible. However, limitations of space and design made it impossible to include every example, suggestion, and explanation within the form itself. We linked frequently within the web form to points in the FAQ addressing specific questions and thematic areas.

The greatest difficulty that we encountered was in formatting the response emails correctly. The Webform module for Drupal does not easily allow for making changes to the output of form submission responses. At this point we required assistance from one of our library web developers. He was able to write a custom script to strip out the question text and format the responses in the

way that we wished researchers to receive them. He also made some changes to the style sheet of the questionnaire itself, which we felt improved the look and feel.

Since the initial implementation, we have made several changes to each component of our DMP online tool to make it more effective. We regularly go back through the informational/instruction page, the questionnaire, and the FAQ to ensure that language/terminology is consistent in all areas. We clarified several questions that had proven confusing to early users of the questionnaire. One example of these changes was to indicate questions that may not be relevant to all researchers, such as HIPAA concerns. As different directorates have provided clarity about their expectations related to DMPs, we have edited the FAQ to reflect this information.

One of our early users also requested the ability to save and return to an in-progress DMP at a later date. This is a complicated process that would require us to hook into the University’s authentication system, which we felt was beyond the scope of our project and the abilities of our content management system. Our University’s authentication system would also have required that the researcher be officially affiliated with the university

**Figure 7. Drupal Webform Module**

Name	Type	Value	Mandatory	E-mail	Operations
+ Introduction, context and basic project information	markup	<!-- #webform-client-form-->			Edit Clone Delete
+ Title of the Project	textfield	-	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Edit Clone Delete
+ Funding Body/Bodies	textfield	-	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Edit Clone Delete
+ Principal Investigator (PI)	textfield	-	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Edit Clone Delete
+ Address of PI	textarea	-	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Edit Clone Delete
+ PI Telephone Number	textfield	-	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Edit Clone Delete
+ PI Email Address	email	-	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Edit Clone Delete
+ For help	markup	<p>For help&nbsp;&filling out t...			Edit Clone Delete
+ End Page 1	pagebreak	-			Edit Clone Delete
+ I. Types of Data	markup	<!-- #webform-client-form-4-->			Edit Clone Delete
+ 1.1 Briefly describe your data, including data types and amount (if known) that will be created or captured	textarea	-	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Edit Clone Delete
+ 1.2 How will you capture or create the data?	textarea	-	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Edit Clone Delete
+ 1.3 Some researchers use existing data in their project. If you are using existing data, where did you get this data and how does it relate to your project?	textarea	-	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Edit Clone Delete
+ For help filling out this form, email mreilly@uh.edu	markup	<p>For help filling out this f...			Edit Clone Delete
+ End Page 2	pagebreak	-			Edit Clone Delete
+ II. Data and Metadata Standards	markup	<!-- #webform-client-form-4-->			Edit Clone Delete
+ 2.1 Which file formats will you use for your data, and why?	textarea	-	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Edit Clone Delete
+ 2.2 To make the data that you capture meaningful what contextual details will be needed? (i.e., Title, Description, Date, etc.)	textarea	-	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Edit Clone Delete
+ 2.3 How do you plan to capture or create metadata for your project? Describe any software or tools that you will use	textarea	-	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Edit Clone Delete
+ 2.4 If you plan to use a specific metadata standard, which one do you plan to use and why?	textarea	-	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Edit Clone Delete
+ For help filling out this form, email mreilly@uh.edu	markup	<p>For help filling out this f...			Edit Clone Delete
+ End Page 3	pagebreak	-			Edit Clone Delete

(it is not possible to create guest network accounts for unaffiliated researchers). This could create difficulties for our researchers working on a collaborative project with colleagues outside of the University of Houston. In its current form, any researcher from any institution can use the form without a login.

## NEXT STEPS

### Local Development and Promotion

Although the multi-institution DMP tools discussed earlier (see the *Tool Review* sidebar) are now widely available, we plan to continue to support and update our local tool. While these other forms are undoubtedly useful and have some valuable features that ours does not (e.g. the ability to save work in progress and link to previous DMPs), we feel that our ability to insert extremely targeted and specific local instructions (e.g. University retention requirements) for our researchers outweighs the absence of those features.

Beyond the ability to provide institution-specific instructions, there are two additional important reasons why we continue to update and provide more context to our form instead of using and promoting one of the multi-institution tools. First, we retain local administration of the form, making it possible for us to follow up with researchers who have used the form and to make decisions about local needs. Second, we can notify researchers if the University of Houston Division Of Research (DOR) or a funding agency makes changes that impact their projects. We are able to do this because our DMP online form keeps a record of every researcher that uses the form. This allows us to offer an expanded level of service and maintain a closer level of involvement with researchers. There may come a time when we move to another form, but for now we feel strongly that our existing form is most beneficial for our researchers. Of course, we certainly do not dissuade our researchers from using other tools if they wish—but we do not actively promote them.

We do, on the other hand, want to further promote the use of our form. There have been 14 total form submissions: nine in 2011, two in 2012, three in 2013. There were eight submissions in the first six months of the form's existence, but then activity dwindled as we focused on updating and improving the form instead of promoting it. Since we have changed our focus back to promoting

the form, there have been five submissions in the past six months. More tellingly, we have seen increased web traffic to the form components on our website, particularly to the printable questions which researchers can use to prepare responses to the form. (Visits to the printable questions peaked sharply in October of 2011 and 2012, indicating a particularly busy time of year for researchers working on funding applications). Overall, there were 49 visits to the printable questions page in 2011, 59 visits in 2012, and 21 visits in the first quarter of 2013, putting us on pace to exceed the previous number in the current year. Although this usage data shows that use is up from the first year of the form's existence, an ongoing initiative in the evolution of our DMP online form is to encourage its use in the wider University community. The DOR guide for researchers links to the form but finding the link is not easy. They are in the process of redesigning their researcher checklist webpage and the DMP online web form will have a more prominent place in the new design. We also hope that our investigation of research support needs (discussed in the following section) will highlight additional opportunities to promote the tool.

### Building Out from the DMP Tool

While researchers who have used the tool have reported that it is extremely helpful, the Head of Digital Services still frequently receives phone calls and emails from researchers. These more personal interactions allow her to provide further explanation of the requirements or to walk the researcher through writing the DMP. Some of the most common questions she receives are: What is metadata? How can I share my data? How long should I preserve my data? What data should I be preserving, working data or completed data? The biggest question of all seems to be: Where can I preserve my data (preferably for free)? These questions have become more complex over time, indicating that our researchers are interested and actively involved in data management planning.

Though our online tool is proving useful for researchers who need to create a DMP, the form (and its instructions and FAQ) is not intended to answer every question about data management. Building on the success of the tool, we have also put together a team of librarians that work in collaboration to provide researchers on campus the best and most current information about data curation and management. One of the team's first projects was to construct a research guide (University of Houston

Libraries, 2013) that explains what data management is, offers best practices, metadata definition and standards, and describes data storage options available on or off-site. This research guide is linked from the DMP online web form and the DMP online web form is linked from the research guide, which provides better information for researchers completing the online form as well as better visibility for the online form with researchers who visit the guide first. A small group, composed of a new librarian dedicated to science research and the DMP form project team, is also currently conducting an investigation of research support needs on campus. This is a follow up to a previous study evaluating researchers' readiness to comply with agency DMP requirements (Peters & Dryden 2011), and it should further inform our efforts to develop services like the DMP tool for researchers.

## RECOMMENDATIONS

A few years have passed since the mandates for submitting DMPs have been put into place. There are now several widely available tools for creating data management plans, and many institutions may prefer to implement an existing tool or otherwise encourage their researchers to use them on their own. However, creating a local tool is still a valuable and achievable option for institutions that would like more local administration without joining or subscribing to an existing tool. The authors recommend building your own tool as an option for institutions that meet the above criteria. A similar tool can be created using most webform utilities.

Building a tool in-house affords the institution the ability to include very specific local information related to their researcher funding activities. This could include resources and services available on campus as well as institution specific policies and requirements, such as Institutional Review Board (IRB) guidelines. An in-house tool allows librarians to monitor form usage and to respond to researcher help requests. Our tool serves an outreach function as well as providing a valuable service to the campus researchers.

## ACKNOWLEDGEMENTS

For more information about the Drupal code modifications, please contact Sean Watkins at [swatkins@uh.edu](mailto:swatkins@uh.edu). We would like to thank Sean for his technical help with this project.

## REFERENCES

- Association of Research Libraries. (2006). To stand the test of time: Long-term stewardship of digital data sets in science and engineering. Retrieved from <http://www.arl.org/bm~doc/digdatartp.pdf>
- Bracke, M. S. (2011). Emerging data curation roles for librarians: A case study of agricultural data. *Journal of Agricultural & Food Information, 12*(1), 65-74. <http://dx.doi.org/10.1080/10496505.2011.539158>
- Brandt, D. S. (2007). Librarians as partners in e-research: Purdue University Libraries promote collaboration. *College & Research Libraries News, 68*(6), 365-396.
- California Digital Library. (2013) DMPTool: Guidance and resources for your data management plan. Retrieved from <https://dmp.cdlib.org/>
- Choudhury, G. S. (2008). Case study in data curation at Johns Hopkins University. *Library Trends, 57*(2), 211-220. <http://dx.doi.org/10.1353/lib.0.0028>
- Choudhury, G. S. (2010). E-science at Johns Hopkins University. In Marcum, D. B. & George, G., (Eds.) *The data deluge: Can libraries cope with e-Science?* Santa Barbara, CA: ABC-CLIO, LLC.
- Cox, A., Verbaan, E., & Sen, B. (2012). Upskilling liaison librarians for research data management. *Ariadne, 70*. Retrieved from <http://www.ariadne.ac.uk/issue70/cox-et-al>
- Cragin, M. H., Palmer, C. L., Carlson, J. R., & Witt, M. (2010). Data sharing, small science and institutional repositories. *Philosophical Transactions of the Royal Society A, 368*(1926), 4023-4038. <http://dx.doi.org/10.1098/rsta.2010.0165>
- Dietrich, D., Adamus, T., Miner, A., & Steinhart, G. (2012). Demystifying the data management requirements of research funders. *Issues in Science & Technology Librarianship, 70*. <http://dx.doi.org/10.5062/F44M92G2>
- Digital Curation Centre (2011). Checklist for a Data Management Plan v3.0. Retrieved from [http://www.dcc.ac.uk/sites/default/files/documents/data-forum/documents/docs/DCC\\_Checklist\\_DMP\\_v3.pdf](http://www.dcc.ac.uk/sites/default/files/documents/data-forum/documents/docs/DCC_Checklist_DMP_v3.pdf)

- Digital Curation Centre. (2013). DMPonline: The DCC data management planning tool. Retrieved from <https://dmponline.dcc.ac.uk/>
- Goldstein, S. & Oelker, S. K. (2011). Planning for data curation in the small liberal arts college environment. *Sci-Tech News*, 65(3), 5-11.
- Harris-Pierce, R. & Liu, Y. Q. (2012). Is data curation education at library and information science schools in North America adequate? *New Library World*, 113(11/12), 598-613. <http://dx.doi.org/10.1108/03074801211282957>
- Integrated Earth Data Applications. (2013). Data management plan tool v.2. Retrieved from <http://www.iedadata.org/compliance/plan>
- Integrated Earth Data Applications. (2013) GeoPass Login. Retrieve from <http://geopass.iedadata.org/josso/>
- Johnston, L. & Jeffryes, J. (2013). Data management skills needed by structural engineering students: A case study at the University of Minnesota. *Journal of Professional Issues n Engineering Education and Practice*, 139(2). [http://dx.doi.org/10.1061/\(ACSE\)EI.1943-5541.0000154](http://dx.doi.org/10.1061/(ACSE)EI.1943-5541.0000154)
- Lage, K., Losoff, B., & Maness, J. (2011). Receptivity to library involvement in scientific data curation: A case study at the University of Colorado Boulder. *portal: Libraries and the Academy*, 11(4), 915-937. <http://dx.doi.org/10.1353/pla.2011.0049>
- Latham, B. & Poe, J. W. (2012). The library as partner in university data curation: A case study in collaboration. *Journal of Web Librarianship*, 6(4), 288-304. <http://dx.doi.org/10.1080/19322909.2012.729429>
- Parham, S. W., Bodnar, J., & Fuchs, S. (2012). Supporting tomorrow's research: Assessing faculty data curation needs at Georgia Tech. *College & Research Libraries News*, 73(1), 10-13. Retrieved from <http://crln.acrl.org/content/73/1/10.full>
- Peters, C., & Dryden, A. R. (2011). Assessing the academic library's role in campus-wide research data management: A first step at the University of Houston. *Science & Technology Libraries*, 30(4), 387-403. <http://dx.doi.org/10.1080/0194262X.2011.626340>
- Scaramozzino, J. M., Ramirez, M. L., & McGaughey, K. J. (2012). A study of faculty data curation behaviors and attitudes at a teaching-centered university. *College & Research Libraries*, 73(4), 349-365. Retrieved from <http://crl.acrl.org/content/73/4/349.full.pdf+html>
- Shorish, Y. (2012). Data curation is for everyone! The case for master's and baccalaureate institutional engagement with data curation. *Journal of Web Librarianship*, 6(4), 263-273. <http://dx.doi.org/10.1080/19322909.2012.729394>
- Steinhart, G. (2007). DataStaR: an institutional approach to research data curation. *IASSIST Quarterly*, 31(3-4), 34-39. Retrieved from <http://www.iassistdata.org/downloads/iqvol313steinhart.pdf>
- Steinhart, G., Chen, E., Arguillas, F., Dietrich, D., & Kramer, S. (2012). Prepared to plan? A snapshot of researcher readiness to address data management planning requirements. *Journal of eScience Librarianship*, 1(2), 63-78. <http://dx.doi.org/10.7191/jeslib.2012.1008>
- Steinhart, G., & Qin, J. (2012). Mentoring for emerging careers in eScience librarianship: An iSchool – academic library partnership. *Journal of eScience Librarianship*, 1(3). <http://dx.doi.org/10.7191/jeslib.2012.1017>
- University of Houston. (2010). Plan to achieve recognition as a National Research University.
- University of Houston Libraries. (2013). Data management. Retrieved from <http://guides.lib.uh.edu/DMP>
- Witt, M. (2008). Institutional Repositories and research data curation in a distributed environment. *Library Trends*, 57(2), 191-201. <http://dx.doi.org/10.1353/lib.0.0029>
- Witt, M., Carlson, J., Brandt, D. S., & Cragin, M. H. (2009). Constructing data curation profiles. *The International Journal of Digital Curation*, 3(4), 93-103. Retrieved from <http://www.ijdc.net/index.php/ijdc/article/view/137/165>

## CORRESPONDING AUTHOR

Michele Reilly  
Head of Digital Services

University of Houston Libraries  
114 University Libraries  
Houston, TX 77204-2000

[mreilly@uh.edu](mailto:mreilly@uh.edu)