

# Getting Everyone Onboard! (and Offboarded): Coordinating Data Services in the Lab

## Authors

Sarah Hauserman, Research Data Management  
*IT Research Computing, Harvard Medical School*  
Julie Goldman, Research Data Services Librarian  
*Countway Library, Harvard Medical School*

## Affiliations

Harvard Longwood Medical Area Research Data Management Working Group



## Why is RDM Important for Onboarding & Offboarding?

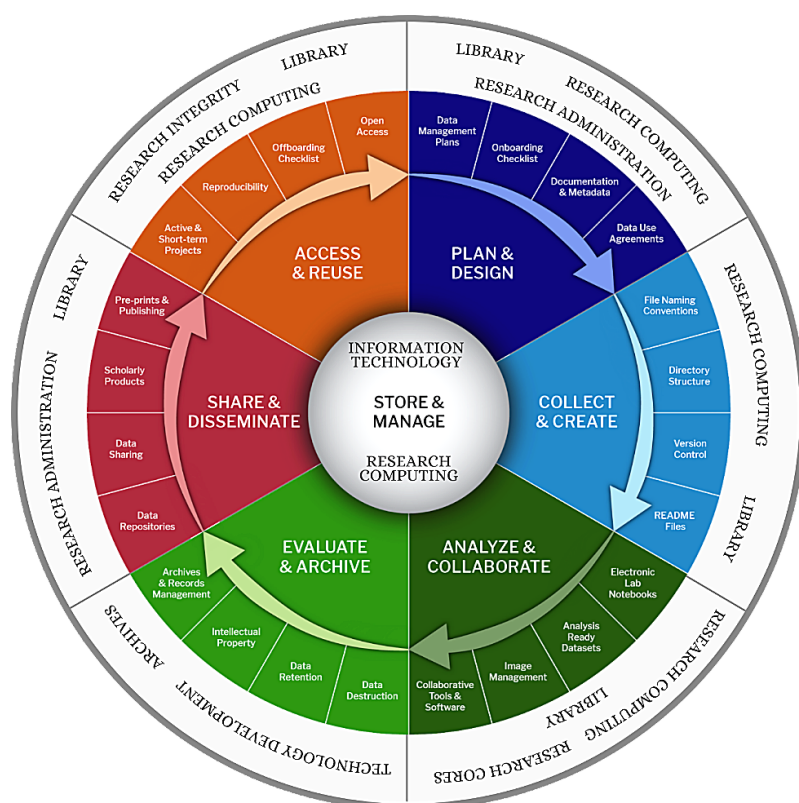
- Traditional employee orientations do not discuss data
- Data accessibility and findability are hindered with improper offboarding
- An employee's departure can negatively impact segments of the data lifecycle
- Creates lab standards for data collection and storage
- Major funding agencies require data documentation at a project's start
- Motivating researchers and PIs is easier with prevailing policies and procedures

## What are the Benefits of Tailored Resources?

- Captures employee institutional (or project) knowledge
- Documents essential information related to projects and datasets
- Enhances research production and consistency
- Helps research projects progress
- Reduces the amount of wasted time spent by others in the lab
- Reduces possible duplication or unnecessary retention of immaterial data
- Reduces overall storage costs
- Facilitates data sharing and reuse
- Increases visibility and notoriety of lab member research

## Why is Collaboration Important?

Coordinated efforts between data professionals in the library, research labs, and information technology departments can reshape institutional data services. Indeed, such collaborative efforts offer a unique opportunity to implement superior data management workflows while also addressing domain-specific challenges.



## Lab Use Case

The George Church Lab (HMS Genetics Department) adapted our checklists to streamline and improve existing lab protocols and procedures. The new checklists were implemented in 2021 to support the lab's mission while also alleviating some of their administrative burden.

**“The transition to the new checklist has been seamless, and feedback from lab members on the new forms has been overwhelmingly positive.”**  
– **Jessica Weber, Lab Data Manager for the Church Lab**

## Related Resources

Resources created by the LMA Research Data Management Working Group discussed throughout this poster are available in [Open Science Framework \(osf.io/pw7ed\)](https://osf.io/pw7ed).

## Introduction

Research Data Management is a collaborative effort requiring the integration and cooperation of various institutional groups, each providing resources and services to researchers, labs, and departments. Such a partnership enables us to develop and maintain data management templates for researchers to use during the employee onboarding and offboarding processes. Additionally, we preemptively generated a template to assist with the creation of data management plans (DMPs), preparing labs for the new data management requirements stipulated by the National Institutes of Health (NIH) beginning in 2023.

# Data Management Plan Template

**The Harvard Data Management Plan (DMP) Template was created in response to the upcoming NIH Data Management and Sharing Policy. The template provides clear guidance to researchers, with recommended policies and procedures, tailored to a Harvard audience. The carefully calculated response from stakeholders across the institution demonstrates a united effort to meet researcher needs and establish consistent messaging.**

Project Details

Collaborators

Write Paper

Research outputs

Download

Finalize / Publish

This plan is based on the "Harvard Medical School Biomedical DMP Template" template provided by Harvard University - (ver: 5, pub: 2021-10-25).

expand all collapse all

0 / 14

+ Data Collection (0 / 2)

Give a brief description of the data, including any existing data or third-party sources that will be used, in each case noting its content, type and coverage. Outline and justify your choice of format and consider the implications of data format and data volumes in terms of storage, backup and access.

What data will you collect or create?

- What type, format and volume of data?
- Do your chosen formats and software enable sharing and long-term access to the data?
- Are there any existing data that you can re-use?

B I

Save

Harvard example answer

Below are sample texts that can be modified (as relevant) to answer the question

<p><b>Data Formats for Harvard Dataserve</b></p>	<p>Immediately after collection, quantitative data will be converted to Stata, SPSS, R, Excel, CSV formats. These formats are fully supported by the Harvard Dataserve, which will perform archival format preservation, metadata extraction, and validity checks. Deposit in these formats will also enable online analysis, variable-level search, data extraction and re-formatting, and other enhanced access capabilities. Documentation will be deposited in PDF/a, or plain-text formats, to ensure long-term accessibility, with any accompanying sound (in WAV), video, or images separate from the documentation deposited as JPEG 2000 files (with lossless compression) or uncompressed TIFF files.</p>
<p><b>General</b></p>	<p>This project will produce primarily scientific data, on an ongoing basis, that will consist of reported publications and intellectual property rights filings. Data types include (1) experimental procedures utilized to obtain data; (2) materials utilized during experimental procedures; (3) methods used for data measurements; (4) data measurements and analysis and DNA and protein sequences; (5) publications in the form of scientific manuscripts; and, (6) patents.</p>
<p><b>General</b></p>	<p>We will be working with [sequencing/mass spectrometry/structure/etc.] data. This data is to be generated [in-house/by a core facility/etc.], and we expect [N] datasets [totaling 100 megabytes/1 terabyte/1 petabyte/etc.]</p>

Guidance

Comments

Harvard

DMPTool

- Harvard Biomedical Data Management: Data Management Plans
- Webinar Video: Let's Talk Data

expand all collapse all

Data format

+ Give support for working with your data: Research Data and Scholarship

- When saving data, it is important to not only consider where you are saving it but also how you are saving.
  - As you shift from saving your data as you work on it to saving it to preserve a long term, something to consider is your file formats. When saving data over the long term, we recommend saving in formats that are as open, lossless, and unencrypted as possible.
  - Open (non-proprietary) formats are those that can be used and implemented by anyone. In practice, this means that files stored in open formats can be opened and used by a variety of proprietary, free, and open-source software tools rather than just a single piece of software. Open, non-encrypted formats are for

## Knowledge Transfer File Template

**These templates embody tailored research data management resources generated through collaborative efforts. We encourage other institutions and groups to utilize our templates to create their own suitable versions.**

Knowledge Transfer File Template		
Name:		
Position/Title:		
Department:		
/lab/Group:		
Dates Affiliated with Lab:		
Date Submitted:		
<p>Overall guidance: In this document and accompanying tables, please document who has access to the data and lab materials that will remain in the lab. Annotate the data appropriately, referencing additional sources, if needed. If specific data files were included as part of a scientific publication, please cite the publication, with a DOI or PubMed ID, if possible. If proprietary data remains inaccessible to collaborators, ensure the data is converted to an open or common file format (e.g. .csv, .txt) and exported to a shared location, prior to your departure.</p>		
<p>Please complete all relevant fields below.</p>		
Category:	Question:	Response:
1. Lab Notebooks, Collaborators, and Data Storage Locations:	<p><i>Properly protecting research data and materials is a fundamental obligation, grounded in the values of stewardship, integrity, and commitments to the providers and sources of the data and reagents. Additional information about data management can be found on the Harvard Biomedical Research Data Management website.</i></p>	
	1.1 Was your work grant funded? If so, provide grant information such as funding agency, grant title, dates, or grant number, if available.	
	1.2 List collaborators who directly contributed to datasets and materials remaining in the lab. Include name and email address(es) for each	
	1.3 Will the project remain ongoing following your departure? If so, who will continue the project and have access to the data? Indicate names and email address(es) for each individual and what data they will have access to.	
	1.4 Did you utilize an electronic lab notebook (ELN) to document the project? If so, provide information about the ELN and who will have	

Harvard Longwood Medical Area Biomedical  
Research Data Management Onboarding Checklist:  
Abridged Version

