• Ca $175 m budget
• + 750 FTE
• + 5000 volumes/week; 250,000 volumes/yr
• Reputation rests on collections (20 m vol.)
• Millions of digital pages
• 10 m photographs
• 1 m maps
• 400 million manuscripts
But What about Data?
Data, it’s not just important, it’s sexy.... Harvard Business School

Data Scientist: The Sexiest Job of the 21st Century

by Thomas H. Davenport and D.J. Patil
Harvard Library Digital Strategy

• **Digitization**
  - Develop a program for Pan-Harvard digitization projects.
  - Update policy guiding vendor partnership for digitization from our collections.
  - Create a comprehensive digital library for our reformatted collections.
  - Create sustainability plan for digital storage costs.

• **Digital Scholarship**
  - Define needs for medium term storage of digital information.
  - Formalize and increase library support for digital scholarship.
  - Formalize and increase library support for data research management.
  - Create a program for publishing via the Harvard Library.
Harvard Guidelines and Policies

At Harvard

For any research project involving data, it is important you understand your rights and responsibilities. The following links will assist you as you consider issues such as intellectual property, security and privacy.

**Office of Vice Provost for Research**
Lists research policies and contacts, as well as a research checklist that will direct you to the relevant policy or to the appropriate office you will need to contact.

**Office for Sponsored Programs**
Lists additional University policies, such as Retention of Research Data and Materials, that correspond to a particular phase of the award life-cycle.

**General Records Schedule** (University Archives)
Regulates how long to keep different types of records, and whether to send records to an archive or destroy them when they are no longer needed; it covers common types of records found across the University in all formats – including electronic records and email.

**Office of Technology Development**
Includes helpful information on licensing, intellectual property, material transfer along with other guidelines, policies, forms and agreements.

Useful Contacts

Research Development (FAS):

**Susan Gomes**, Director of Research Development and Strategy
phone: (617) 496-9448

**Erin Cromack**, Research Development Officer
phone: (617) 496-5252

Sponsored Research Administration, Social Sciences (FAS):

**Pamela Baker-Webber**, Sponsored Research Administrator
phone: (857) 488-6890

Harvard Medical School Research Administration

Sponsored Research Administration (HMD/HSDM)

Research Records Management (University Archives)
Serves Central Administration, the graduate and professional schools (other than those listed below), research
Data Symposium

Harvard Purdue Data Management Symposium

June 2015

Symposium Home
Joint Introduction
Harvard Policy for Data
Research Data Management: Our Collaborative Future
The Research Data Revolution
The Researcher’s Perspective
Internal Drivers: Researchers and Data
External Drivers: Funders and Publishers
Repository Services

HARVARD UNIVERSITY LIBRARIES

PURDUE UNIVERSITY LIBRARIES

hpdm2015
@HarvardLibrary
@PurdueLibraries

This symposium was generously sponsored by Arcadia and Purdue University.

JUNE 16-17, 2015

This symposium on data management incorporated visionary ideas, new concepts, and inspirational speakers.
Data Scientist Training for Librarians course has created many valuable learning outcomes that can help anyone interested in becoming a data scientist/data librarian. Tools for analyzing and cleaning data, software for visualizing data and more at http://altbibli.io/dst4l/blog/
Launch of Data Management Planning Tool

April 26, 2016
Author: Eleni Castro

As a result of collaborations with the Office of the Vice Provost for Research, Harvard University Information Technology, and IQSS, Harvard Library has launched a customized version of DMPTool, an online data management planning tool, for Harvard University. Data management plans—documents that outline what researchers will do with data during and after a project—are becoming increasingly required by funding agencies such as the National Institutes of Health and the National Science Foundation. The online tool provides step-by-step guidance for
DMPT

Data Management

What is DMPTool?
DMPTool is an online tool available to help you create and share your data management plans to meet funder requirements and as a best practice for managing your data. DMPTool provides step-by-step guidance for creating your own DMP and includes templates and sample plans to help you address requirements specific to Harvard and your funding sources.

Using DMPTool at Harvard
Harvard is a DMPTool partner institution. Log in using your Harvard credentials.

- Go to [DMPTool](#). Click Get Started.
- Select Harvard University from the menu of DMPTool partner institutions, then click Next.
- Enter your Harvard credentials when prompted.

Benefits of Using DMPTool (video)

[DMPTool2 Promotional Video](#) from California Digital Library.
Data Management

Why manage data?

- **Organize your research so it can be reproduced.** Managing your data makes it easier to understand the details and procedures relating to your data and data collection throughout the lifecycle of the project.

- **Preserve and share your data to get recognition.** The data you collect are the basis of your research. They are your unique contribution, and preserving them means that your work will be recognized by others. It also ensures that your work can support future research and facilitate new discoveries.

- **Satisfy funding requirements and Harvard research policies.** The number of granting bodies requiring that data be preserved and shared is growing. A good data management plan will help you meet the requirements of your funding agency and address preservation, documentation, and verification issues. It helps reviewers understand the characteristics of your data.

What is a data management plan?

A **data management plan**, or DMP (sometimes also called a data sharing plan), is a formal document that outlines what you will do with your data during and after a research project. Most researchers collect data with some form of plan in mind, but it's often inadequately documented and incomplete. Many data management issues can be handled easily or avoided entirely by planning ahead. With the right process and framework, it doesn't take too long and can pay off enormously in the long run.

Many funding agencies, especially government funding sources, require a DMP as part of their application processes. Even if you are not seeking funding for your research, documenting a plan for your data is a best practice and will help your data comply with Harvard's policies.
Welcome!

Particularly when we talk about large numbers, it can be difficult to fully understand their impact. With an ever-increasing amount of data and information available to us, data visualization is becoming more important to help people truly understand the meaning of the information that is collected. Whether you are teaching in a classroom or presenting in front of clients, the ability to distill and contextualize data is one that will set you apart and the tools in this guide will help you to do just that.

If you are new to the world of data visualization and want an overview, try...
STORYTELLING WITH DATA WORKSHOPS AT BOSTON UNIVERSITY

JUNE 6-10
JUNE 13-17
At Research Data Services, researchers across disciplines benefit from expert consultation and training in acquiring, collecting, wrangling, analyzing, visualizing, sharing, and preserving research data. Our data professionals, in collaboration with our research librarians, work to advance data science, using data to answer scientific questions, making analysis reproducible, and promoting open data to enable long-term discovery and knowledge.

The Social, Natural, & Engineering Science team supports the research needs across scientific disciplines, from finding materials and managing information, to understanding the changing scholarly landscape and promoting information literacy. Information experts, together with our data team, work with students and faculty to enhance their scholarship and navigate resources across the University.
MANAGING YOUR DATA

Got data? We're here to help you manage, share, and preserve your research data. In addition to our Data Repository for the U of M curation services, the Libraries will help you navigate available campus resources throughout the data lifecycle:

**Before Your Research Begins**

- Schedule a data management plan (DMP) consultation ([Request Form](#)) or use our [DMP templates](#).
- Explore [funding agency requirements](#) for data and learn best practices for getting [IRB approval](#) for sharing data.
- See [more tools for planning for data management](#).

**During Your Research**

Need help? Contact Us with your questions and we will consult with you or point you to the right person, resource, or service on campus.
Increasing Openness and Reproducibility in Quantitative Research
A Workshop by the Center for Open Science

Tuesday, May 3, 2016 - 9:00am to 12:00pm
Library Conference Center

PLEASE NOTE VENUE CHANGE: The workshop will be held in the Main Conference Room (not the WEC) in the Research Library. Signs will be posted to direct attendees.
<table>
<thead>
<tr>
<th>Name</th>
<th>Position</th>
<th>Expertise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stephen Balogh</td>
<td>Data Services Specialist</td>
<td>QGIS, ArcGIS, Ruby, JavaScript</td>
</tr>
<tr>
<td>Andrew Battista</td>
<td>Librarian for Geospatial Information Systems</td>
<td>Google Fusion Tables, Google Earth, ArcGIS</td>
</tr>
<tr>
<td>Sarah DeMott</td>
<td>Qualitative Data Analysis &amp; Survey Design Specialist</td>
<td>ATLAS.ti, Qualtrics, NVivo</td>
</tr>
<tr>
<td>Denis Rubin</td>
<td>Senior Academic Technology Specialist (Quantitative Services)</td>
<td>MATLAB, SPSS, SAS, Getting Started with R, Qualtrics, Python</td>
</tr>
<tr>
<td>Matthew Sumner</td>
<td>NYU Abu Dhabi Data Services Librarian</td>
<td>ArcGIS, Google Earth, Google Fusion Tables</td>
</tr>
<tr>
<td>Michelle Thompson</td>
<td>Geographic Information Systems Specialist</td>
<td>ArcGIS, Google Earth, Google Fusion Tables</td>
</tr>
</tbody>
</table>
Is Research Reproducibility the New Data Management for Libraries?

Cynthia R.H. Vitale

Version of Record online: 24 FEB 2016
DOI: 10.1002/bul2.2016.1720420313

Copyright © 2016 American Society for Information Science and Technology

Bulletin of the Association for Information Science and Technology

Research Data Roadshow

Call for Two Curriculum Designer/Presenters

for Research Data Management (RDM) Roadshow

ACRL is accepting applications from prospective curriculum designer/presenters to create professional development on Research Data Management. A new, day-long workshop will provide professional development support to librarians who wish to know more about data management. The target audience for this workshop is library administrators, subject liaison librarians, and other non-data specialists. Online content will be created to complement the roadshow and may include stand alone webinars.

The curriculum designers will plan and develop learning experiences for face to face and virtual learning intended to build librarians’ capacity as leaders on issues/projects of campus-wide interest. Potential topics could include:

- the role of data information literacy vis-à-vis scholarly communication
- exploring the role of the liaison within the data management conversation
- how to assist with data management plans
- organization, documentation, storage, and security best practices
- discipline-specific data concerns
  - metadata standards
  - finding data
- administrative strategies
  - defining scope, setting vision
  - creating partnerships
  - resourcing data management
  - creating a program statement for the role or set of roles responsible for data management

As they develop the curriculum, a small number of ACRL member leaders, including presenters for other related workshops, will act as advisors. The curriculum designer/presenters will build on the experience of a preconference workshop offered at ACRL 2015 on research data management.
M-13-13

MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

FROM: Sylvia M. Burwell
Director

Steven VanRoekel
Federal Chief Information Officer

Todd Park
U.S. Chief Technology Officer

Dominic J. Mancini
Acting Administrator, Office of Information and Regulatory Affairs

SUBJECT: Open Data Policy—Managing Information as an Asset

Information is a valuable national resource and a strategic asset to the Federal Government, its partners, and the public. In order to ensure that the Federal Government is taking full advantage of its information resources, executive departments and agencies (hereafter referred to as “agencies”) must manage information as an asset throughout its life cycle to promote openness and interoperability, and properly safeguard systems and information. Managing government information as an asset will increase operational efficiencies, reduce costs, improve services, support mission needs, safeguard personal information, and increase public access to valuable government information.
Open Government Data (2016)
Research Article

Approaches to Data Sharing: An Analysis of NSF Data Management Plans from a Large Research University

Authors: Carolyn Bishoff, Lisa Johnston

Abstract

INTRODUCTION Sharing digital research data is increasingly common, propelled by funding requirements, journal publishers, local campus policies, or community-driven expectations of more collaborative and interdisciplinary research environments. However, it is not well understood how researchers are addressing these expectations...
What’s the future?

• Expand staffing with research data skills
• Partner with others: Research, IT, faculty
• Think of digital libraries as data
• Create easy-to-find integrated and coordinated services