One academic library's response to the data dilemma

Courtney Crummett

MIT Libraries

Bioinformatics and Biosciences Librarian
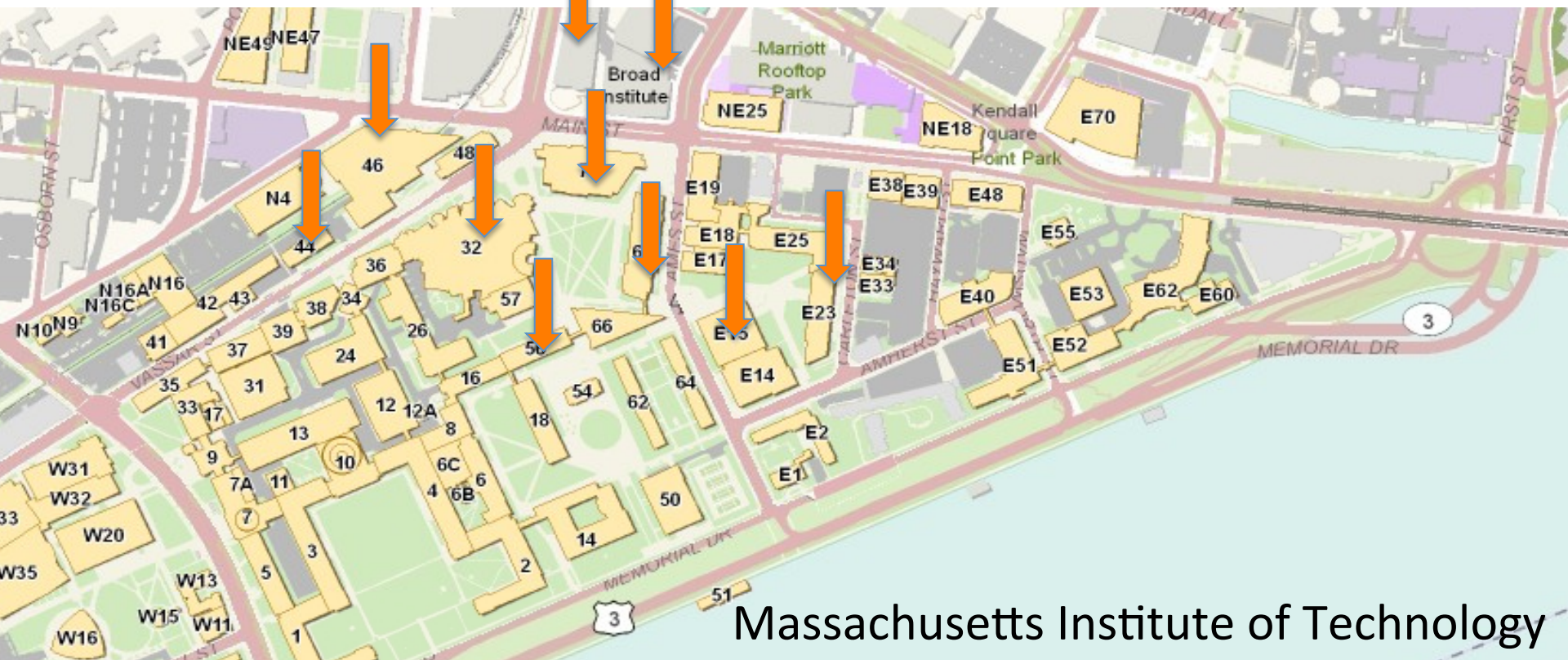MIT Libraries Research Data Services Working Group Member

MLA 2013
May 6, 2013

MITLibraries

Cambridge

Massachusetts Institute of Technology

Boston
You are located down this street.

MITLibraries

Biology

Brain & Cognitive Science

Biophysics

Computational & Systems Biology

Microbiology

Molecular Neuroscience

Health Sciences and Technology

Biological Engineering

Koch Institute for Integrative Cancer Research

Institute for Medical Engineering & Science

Picower Institute for Learning & Memory

McGovern Institute for Brain Research

Whitehead Institute for Biomedical Research

Broad Institute of Harvard and MIT
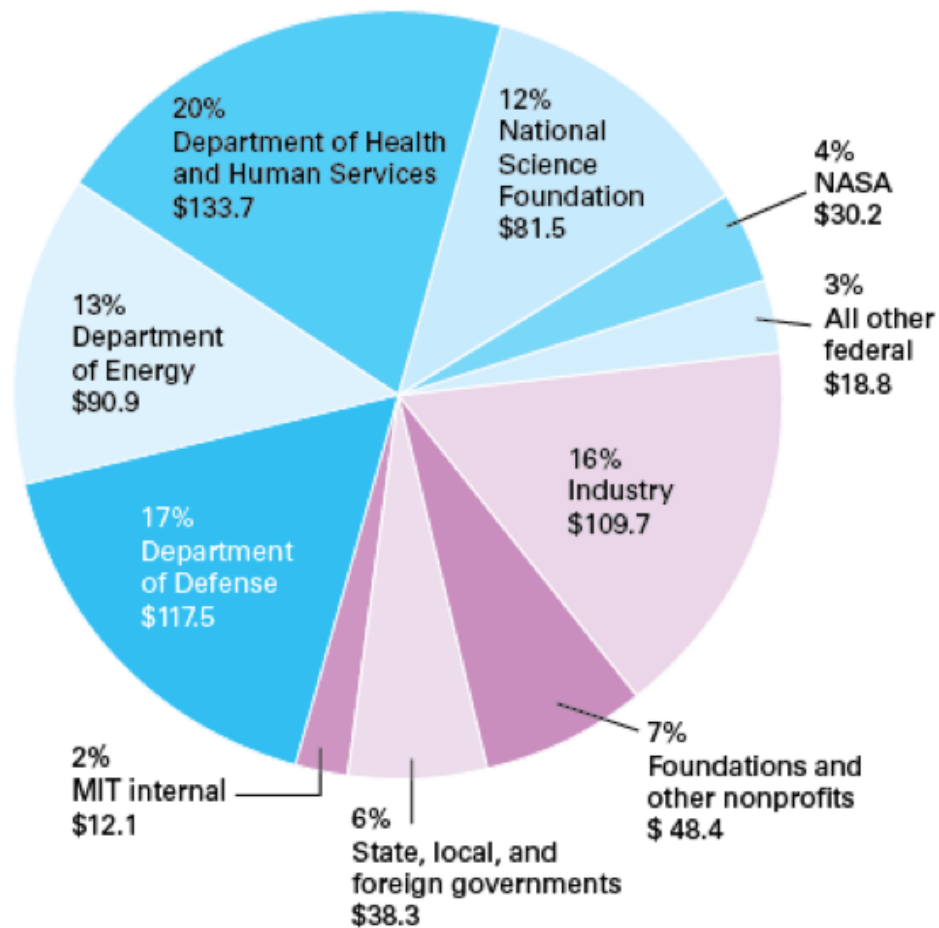
Biomedical Research Milestones at MIT

- Large contributions to the Human Genome Project
- Discovery of genetic basis of Lou Gehrig's disease and Huntington's disease
- 10 Nobel Prizes in Medicine/Physiology
  - Horvitz 2002 for discovery and characterization of key genes controlling cell death
- Grew Vacanti "earmouse" with U Mass Medical

MITLibraries

**Fiscal Year 2012
Research Expenditures by
Primary Sponsor
(figures in millions)**

Total: $681.1 million

- Research
- Research Expenditures
- Interdisciplinary Centers, Labs, & Programs
- Research Alliances

20%
Department of Health and Human Services
$133.7

12%
National Science Foundation
$81.5

4%
NASA
$30.2

3%
All other federal
$18.8

13%
Department of Energy
$90.9

16%
Industry
$109.7

17%
Department of Defense
$117.5

2%
MIT internal
$12.1

6%
State, local, and foreign governments
$38.3

7%
Foundations and other nonprofits
$48.4

Research Data Management Services at MIT Libraries

History

2005 - Data Initiatives Group (DIG)

2008 - Data Management and Publishing guide

2011 - Research Data Management Team formed with primary objectives of outreach, education, resources and collaboration.

    Established Framework of Services
    ARL/DLF E-Science Institute

2012 – Research Data Services working group

MITLibraries

Current Research Data Services

Answer data-management@mit.edu

Maintain the MIT Libraries' web presence for research data services

Maintain the data management guide

Assist with creating data management plans

Develop and provide workshops

Consult and collaborate with researchers

Refer to related services within the Libraries or Institute

Manage Your Data

The MIT Libraries supports the MIT community in the management and curation of research data by providing the following services:

Data Management Guide

This Data Management and Publishing Guide is a practical self-help guide to the management and curation of research data throughout its life cycle. It provides guidance on a range of topics, including: planning for data management, documentation/metadata, file formats, data organization, data security and backup, citing data, data integration, funder requirements, ethical and legal issues, and sharing and archiving data.

Assistance with Creating Data Management Plans

Many funders, such as the National Science Foundation, have requirements for data sharing and data management plans. We can help you to put together such a plan, assess the data management needs of your particular project, and assist in identifying solutions for data management and archiving.

Workshops

Managing Research Data 101, a workshop designed for researchers working to manage their data, offers basic, practical strategies for data management, providing an overview of the topics covered in the Data Management and Publishing Guide. The latest version of the slides from this workshop can be found at Managing Research Data 101 (PDF).

Individual Consultation and Collaboration with Researchers

We are available for individual consultation on data management issues, and can provide expertise in areas such as data organization and preservation, connect you to a network of data management services, and advocate for your needs. We can help you in understanding your data management needs and recommending optimal practices for keeping your data usable, now and into the future.

Referrals to Related Services

DSpace@MIT is a possible option for making your data publicly available and ensuring long term access to it, depending on the nature of the content type and the size of the datasets. See the Guidelines for Research Dataset Contributions in DSpace@MIT for more details. Beyond MIT, we can assist you in finding appropriate subject based data repositories for your work.

Contact Us

For help with your data management needs please contact us at data-management@mit.edu. The Libraries' Research Data Management Services are an evolving set of services, and we welcome your feedback on them as well as your input on what kinds of data management services you would find most valuable.

http://libraries.mit.edu/guides/subjects/data-management

Research Data Services Workshops

Managing Research Data 101 Workshop

Research Data 102 Workshops:
Version Control
File Management (in development)

Research Data Management Panel

Case Study #1 PhysioNet

Researcher: PI from Health Sciences and Technology Division

Type of data: repository of physiological data and software

Data management issue: long term curation and preservation of repository and support to user community

Case Study #1 PhysioNet Continued

Challenges:

- unique request
- not just MIT data; not just MIT users
- support requires a strong subject background

Service we provided:

- consultation and advice
- exploring alternative solutions
- advocacy

MITLibraries

Case Study #1 PhysioNet Continued

Outcomes:

- Learned more about data management issues
- Relationship building
- Learned about physiological data
- Increased communication with PI
- Exploring possible library pilot project

MITLibraries

Case Study #2: NIH Grants Policy Statement

Researcher: Koch Institute PI

Type of data:
- Cell processes and immunology
- raw & quality controlled images
- figures
- manuscripts

Data management issue: proper archive solution for storing data for 3 years past the grant closing date in accordance with the "NIH Grants Policy Statement."

MITLibraries

Case Study #2: NIH Grants Policy Statement Cont'd

Service provided:
- Consultation and advice
- Exploring alternative solutions: DSpace@MIT, DataVerse, or Massachusetts Green High Performance Computing Center (MGHPCC)

Outcomes: (more of the same)
- Learned more about data management issues
- Relationship building
- Increased communication with PI
- Awareness about solutions (MGHPCC)
- Discovery of future services to offer PI

MITLibraries

NSF data management plan consultations

Number of requests: 3

Faculty from NSF Directorates:
- Geological Sciences
- Civil, Mechanical and Manufacturing Innovation (CMMI) – Civil Infrastructure
- Civil, Mechanical and Manufacturing Innovation: Structural Materials and Mechanics

Staff time: 35 hours

Data: images, electronic microscopy images, programming code, numerical data

MITLibraries

Other mentionables

Helped Economics Faculty with data documentation for data deposit to the Inter-university Consortium for Political and Social Research (ICSPR); advocated on behalf of faculty to ICSPR regarding deposit; ongoing help with documentation.

Advised Research Scientist on selecting databases for storing and sharing data and discussed appropriate file metadata for file management.

Opportunities and Barriers

- Research priority

- Getting the word out

- MIT culture

- Relationships with other offices on campus

- Solutions are still new

MITLibraries

One academic library's response to the data dilemma

- We have a team in place offering various services to the community
- Consultations are productive and learning goes both ways
- We have had very few NSF data management plan requests
- Data management isn't a priority (yet)
- Advising is as important as providing a solution

MITLibraries