Data Management Plans & the DMPTool
Data Management Services @ MIT Libraries

• Workshops/Webinars
• Web guide: http://libraries.mit.edu/data-management
• Individual consultations
  ▪ includes help with creating/reviewing data management plans
Why a DMP? Why data management?

- Meet funder requirements
  - DMPs required within grant applications
  - Researchers required to share data collected
- Saves time and prevents data loss
- Makes it easy to share data later, increasing research impact
“The Obama Administration is committed to the proposition that citizens deserve easy access to the results of scientific research their tax dollars have paid for.”
Funder Requirements

**Research Funder Open Access Requirements**

The table below summarizes major US Research funder open access requirements for publications and data, and links to information on related journal requirements.

For questions, please contact:

- For publications: Ellen Finnie, Head, Scholarly Communications & Collections Strategy
- For data: data-management@mit.edu
- See also more information on the Research Funder Policies and Related Legislation page.

List of major funder requirements:

http://libraries.mit.edu/scholarly/publishing/research-funders/research-funder-open-access-requirements/

<table>
<thead>
<tr>
<th>Agency / Entity</th>
<th>Publications</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>White House Office of Science and</td>
<td>All federal agencies over $100M annually in R&amp;D required to make peer-reviewed</td>
<td>All federal agencies over $100M annually in R&amp;D required to make</td>
</tr>
<tr>
<td>Technology Policy Directive</td>
<td>manuscripts available within 12 months. 2013, with implementation beginning</td>
<td>data scored and publicly accessible. Scope: data necessary to validate</td>
</tr>
<tr>
<td>Office of Science &amp; Technology Policy</td>
<td>October 2014 and timetable varying by agency. View the Directive</td>
<td>research findings, including data sets used to support scholarly</td>
</tr>
<tr>
<td></td>
<td></td>
<td>publications. (Lab notebooks not included)</td>
</tr>
<tr>
<td>DoD</td>
<td>Public Access Plan Draft in 2015, takes effect FY16</td>
<td>Public Access Plan Draft in 2015, compliance begins FY17</td>
</tr>
<tr>
<td>US Department of Defense</td>
<td>Authors will be required to upload their final accepted peer-reviewed journal</td>
<td>Data Management Plans (DMPs) will be required, focusing on data</td>
</tr>
<tr>
<td></td>
<td>manuscripts (or the final published article, if the author has sufficient</td>
<td>&quot;necessary to validate research findings.&quot;</td>
</tr>
<tr>
<td></td>
<td>rights) to the Defense Technical Information Center (DTIC) at the time of</td>
<td>Metadata for each data set, &quot;including subject, characteristics,</td>
</tr>
<tr>
<td></td>
<td>acceptance. Articles will be available no later than 12 months following</td>
<td>and location, will be shared via DTIC’s DoD data set catalog.</td>
</tr>
<tr>
<td></td>
<td>publication.</td>
<td></td>
</tr>
<tr>
<td>DOE</td>
<td>Public Access Plan</td>
<td>Public Access Plan</td>
</tr>
<tr>
<td>US Department of Science</td>
<td>Public Access Plan Draft in 2015</td>
<td>Public Access Plan</td>
</tr>
<tr>
<td>Still Pursuing Research</td>
<td>Draft in 2015, compliance begins FY17</td>
<td>Public Access Plan</td>
</tr>
<tr>
<td>US Department of Science</td>
<td>Draft in 2015, compliance begins FY17</td>
<td>Public Access Plan</td>
</tr>
</tbody>
</table>
Today’s Workshop

1. The What
   Content of a data management plan

2. The How
   The DMPTool
General content of a DMP

1. Project, experiment, and data description
2. Documentation, organization, and storage
3. Access, sharing, and re-use
4. Archiving

Always check for and follow a funder’s specific requirements

Details for each of these areas can be found on our website at: http://libraries.mit.edu/data-management/plan/write/
1. Project, experiments, and data generated

Data description

- What is it?
- How will it be collected?
- What format is it in?
- How much data will be generated?

Secondary data

- Are you using data that someone else produced? If so, where is it from?

Roles and Responsibilities

- Who is responsible for managing the data?
2. Documentation, organization, and storage

Documentation (aka metadata)

- What documentation will you create in order to make the data understandable by other researchers?
  - What types of information are important for describing, discovering & using the data?
  - Which mechanisms for metadata capture make sense given your research workflow?
- Are you using documentation standard to your field?
  - See Digital Curation Centre's browsable database of disciplinary metadata for examples

More information on documentation/metadata can be found on our website at: http://libraries.mit.edu/data-management/store/documentation/
2. Documentation, organization, and storage

Organization

- What directory and file naming conventions will you use?

More information on file organization can be found on our website at: http://libraries.mit.edu/data-management/store/organize/
2. Documentation, organization, and storage

Storage

- What are your local storage and backup procedures?
- Will this data require secure storage?
- What tools or software are required to read or view the data?

More information on backups & MIT resources can be found on our website at: http://libraries.mit.edu/data-management/store/backups/
3. Access, sharing, and re-use

- What data will be shared, when, and how?
  - Will this research be published in a journal that requires underlying data to accompany articles?
  - Will there be any embargoes on the data?

- Does sharing the data raise privacy, ethical, or confidentiality concerns?
  - Do you have a plan to protect or anonymize data, if needed?

More information on sharing data can be found on our website at: http://libraries.mit.edu/data-management/share/
3. Access, sharing, and re-use

- Who holds *intellectual property rights* for the data and other information created by the project?
  - Will any copyrighted or licensed material be used?
  - Do you have permission to use/disseminate this material?

- Are there any *patent- or technology-licensing-related restrictions* on data sharing associated with this grant?

- Will you *permit re-use*, redistribution, or the creation of new tools, services, data sets, or products?
4. Archiving

- Will you store the data in an archive or repository for long-term access? If not, how will you preserve access to the data?
  - Is a discipline-specific repository available?
  - Or is there a funder-mandated repository?

Resources for finding a repository

- Registry for Research Data Repositories: [re3data.org](http://re3data.org)
4. Archiving

- How will you prepare data for preservation or data sharing?
  - Will the data need to be anonymized or converted to more stable file formats?
  - Are software or tools needed to use the data? Will these be archived?

- How long should the data be retained? 3-5 years, 10 years, or ‘forever’?
Today’s Workshop

1. The What?
   Content of a data management plan

2. The How?
   The DMPTool
The DMPTool

Data Management Planning Tool
Create, review, and share data management plans that meet institutional and funder requirements.

https://dmptool.org