Writing a Data Management Plan

From Funding Proposal to Full DMP

Jenny O’Neill, Data Manager, UCD Library
jenny.oneill@ucd.ie
Before we Begin

Data Sharing and Management Snafu in 3 Short Acts

- NYU Health Sciences Library
- https://youtu.be/N2zK3sAtr-4
Overview

• Some definitions & concepts
• Ten rules for creating a good DMP
• Brief summary
• Questions


Syrotiuk, Nicholas (2019): “Writing a good Data Management Plan: a workshop for research active staff & PGRs” DOI: http://doi.org/ddn5
‘research data’ means documents in a digital form, other than scientific publications, which are collected or produced in the course of scientific research activities and are used as evidence in the research process, or are commonly accepted in the research community as necessary to validate research findings and results;
‘research data’ means documents in a digital form, other than scientific publications, which are collected or produced in the course of scientific research activities and are used as evidence in the research process, or are commonly accepted in the research community as necessary to validate research findings and results;
What is data management?

- create
- document
- use
- store
- share
- preserve

Long-term management of data

Day-to-day management of data

Adapted from Activities Involved in RDM, (c) Stuart Macdonald/EDINA. Used with permission.
FAIR Data Principles

- Findable
- Accessible
- Interoperable
- Reusable

Steps towards FAIRer data

• Make data available in a data repository or archive
  – A persistent identifier (PID)
  – Rich metadata (elements which describe the data)
  – A clear licence
Data Management Plans

• A communication tool
  – Maximise re-use potential of data
  – Keep personal, sensitive & confidential data safe
  – Preserve research data for longer
Welcome

DMPonline helps you to create, review, and share data management plans that meet institutional and funder requirements. It is provided by the Digital Curation Centre (DCC).

Join the growing international community that have adopted DMPonline:

17,622 Users
203 Organisations
23,083 Plans
89 Countries

Some funders mandate the use of DMPonline, while others point to it as a useful option. You can download funder templates without logging in, but the tool provides tailored guidance and example answers from the DCC and many research organisations. Why not sign up for an account and try it out?
Sample DMPs

Digital Curation Centre (DCC) Example DMPs and Guidance

• http://www.dcc.ac.uk/resources/data-management-plans/guidance-examples

DMPOnline Public DMPs

• https://dmponline.dcc.ac.uk/public_plans

LIBER Data Management Plan Catalogue

• https://libereurope.eu/dmpcatalogue/
1: DETERMINE FUNDER REQUIREMENTS
Irish Funders

• HRB Policy
  – The HRB requires a data management plan (DMP) supplemental to all approved grant proposals where DMPs requirements are specified within the award call guidance.

• HRB leading the way but IRC and SFI following closely behind.
2: IDENTIFY THE DATA TO BE COLLECTED

- Types of data
- Methods of data collection
- Volume of data in bytes
- File formats: open preferred
• Mention file naming conventions, version control & folder structures
  – Consistent
  – Meaningful to you and your colleagues
  – Easier to find, understand, and re-use

3: DEFINE HOW THE DATA WILL BE ORGANIZED
4: EXPLAIN HOW THE DATA WILL BE DOCUMENTED
Documentation

• Good quality documentation allows others to find and understand your data:
  – Interview protocol
  – Questionnaires & interviewer instructions
  – Codebook or data dictionary
  – Information sheets, Consent forms, Ethical approval
  – Database schemas
  – Methodology reports
  – Provenance information about sources of derived or digitised data

https://mantra.edina.ac.uk/documentation_metadata_citation/
5: DESCRIBE HOW DATA QUALITY WILL BE ASSURED

• Quality assurance
• Quality control
6: PROVIDE A STRATEGY FOR SHORT-TERM DATA STORAGE AND LONG-TERM DATA PRESERVATION

• How long will the data be accessible?
• How will data be stored and protected over the duration of the project?
• How will data be preserved and made available for future use?
<table>
<thead>
<tr>
<th><strong>LEGAL DATA TYPES</strong></th>
<th><strong>AMBER DATA</strong></th>
<th><strong>GREEN DATA</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Protection of data is required by law or regulatory instrument.</td>
<td>UCD has an obligation to protect the data.</td>
<td>Protection of data is at the discretion of the owner or custodian.</td>
</tr>
<tr>
<td>Disclosure would cause exceptional or long term damage to the reputation of the University, or risk to those whose information is disclosed.</td>
<td>Could cause harm to the reputation of the University.</td>
<td>Low risk of embarrassment or reputational harm.</td>
</tr>
<tr>
<td><strong>REPUTATION OF UCD DATA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May have serious or long term negative financial impact on the University.</td>
<td>May have short term financial impact on the university.</td>
<td>No impact to the commercial operation of UCD.</td>
</tr>
<tr>
<td><strong>COMMERCIAL DATA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information which provides access to resources, physical or virtual.</td>
<td>Smaller subsets of protected data from a school.</td>
<td>General university information.</td>
</tr>
<tr>
<td><strong>OTHER INSTITUTIONAL RISKS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>APPROVED STORAGE PLATFORM OPTIONS</strong></td>
<td>Novell Storage. Encrypted Storage Device*</td>
<td>Google Drive Novell Storage Microsoft office 365 Encrypted Storage Device *</td>
</tr>
<tr>
<td></td>
<td>Google Drive Novell Storage Microsoft office 365 Storage Device *</td>
<td>Google Drive Novell Storage Microsoft office 365 Storage Device *</td>
</tr>
</tbody>
</table>
Backup Strategy

3. total copies of your data
2. different mediums (devices)
1. copy offsite
7: DEFINE THE PROJECT’S DATA POLICIES

- licensing or sharing arrangements that pertain to the use of pre-existing materials;
- plans for retaining, licensing, sharing, and embargoing data, code, and other materials;
- legal and ethical restrictions on access and use of human subject and other sensitive data.
Relevant Policies

- EU General Data Protection Regulation (GDPR)
- National policy statement on ensuring research integrity in Ireland [pdf]
- UCD Code of Good Practice in Research
- UCD Data Protection Policy
- UCD Intellectual Property Policy
- UCD Password Protection Policy
- UCD Research Ethics Policy
- UCD Research Integrity Policy
• Choose a discipline specific repository
• If no disciplinary repository exists for your discipline consider depositing data in a multidisciplinary repository

8: DESCRIBE HOW THE DATA WILL BE DISSEMINATED
9: ASSIGN ROLES AND RESPONSIBILITIES
10: PREPARE A REALISTIC BUDGET

- People
- Storage & Computation
- Creation & reuse of data
- Deposition & Preservation
Imagine...

Three years after having completed your graduate degree, a researcher in your field calls to verify the integrity of the data you collected for your dissertation research because she thinks it may be flawed.

- Can you locate your dissertation files?
- Was it stored properly in a location you can still access?
- Are any of the files corrupted? Are they in usable and sharable formats?
- Do you have proper documentation about how you created, analyzed and made use of your data that still makes sense to you and other researchers?
ORGANISE – If you make a plan, you just might follow it

DOCUMENT – What would my colleagues need to know to understand this data?

STORE – 3 copies in 3 locations

SHARE – Data makes an impact (to you, to your research group, to society)

To recap
UCD Library Resources

Research Data Management LibGuide

FAIR Data LibGuide

UCD Library Data Management Checklist

Addressing the FAIR Data Principles in a Data Management Plan

Where to Submit Data
THANK YOU

jenny.oneill@ucd.ie

http://libguides.ucd.ie/data