Loughborough University

Data Management as Part of a Research Workflow

Gareth Cole, Loughborough University

Who am I?

- Research Data Manager at Loughborough since March 2015
- Data Curation Officer at University of Exeter between October 2011 and February 2015
- PhD Maritime History (Exeter) – 2008: “The Office of Ordnance and the arming of the Fleet in the French Revolutionary and Napoleonic Wars, 1793-1815”
- @DrGarethCole
- ORCID: https://orcid.org/0000-0001-7493-0137

What is/are research data?

- **Research data** are the evidence that underpins the answer to the research question, and can be used to validate findings regardless of its form (e.g. print, digital, or physical). These might be quantitative information or qualitative statements collected by researchers in the course of their work by experimentation, observation, modelling, interview or other methods, or information derived from existing evidence. Data may be raw or primary (e.g. direct from measurement or collection) or derived from primary data for subsequent analysis or interpretation (e.g. cleaned up or as an extract from a larger data set), or derived from existing sources where the rights may be held by others. Data may be defined as ‘relational’ or ‘functional’ components of research, thus signalling that their identification and value lies in whether and how researchers use them as evidence for claims.
- **They may include**, for example, statistics, collections of digital images, sound recordings, transcripts of interviews, survey data and fieldwork observations with appropriate annotations, an interpretation, an artwork, archives, found objects, published texts or a manuscript.

Data Management Plan / Technical Plan

“Plans typically state what data will be created and how, and outline the plans for sharing and preservation, noting what is appropriate given the nature of the data and any restrictions that may need to be applied.”

Why is data management important?

Why have a plan 1

- Research efficiency
- Aids collaboration and sharing
- Try to anticipate problems before they arise

Why have a plan 2

- You have an “Intellectual Plan” why not a “Technical Plan”?
- Ensure you have permission to do everything you want to do
- Ensure you have enough resource to do everything you want to do
- Plans should be dynamic – your research changes, technologies/solutions change

The sticks…

- Funder requirements
  - As part of grant application
  - Once grant has been approved
  - Different funders have different requirements
- Publisher requirements

Areas to think about in a plan 1

- Use of existing data
  - “Do I need to create/collect new data?”
- Storage
- Backup
- Ethics, legal, and licences

Areas to think about in a plan 2

- Version control
- Quality control
- Collaborative work – sharing during project
- Responsibilities
- Resources and costs

General hints and tips

- Be realistic
- Cover all the data/information you will collect and create
- Be as specific as possible
- Justify any resources needed
- Don’t overthink it

Tools and support available

- DMPonline – UK
- DMPtool – US
- Mantra – University of Edinburgh
- RDM Mooc
- Local institutional support
- DCC guidance
- DCC DMP checklist - http://www.dcc.ac.uk/resources/data-management-plans/checklist

Write your own DMP

- Use DMPonline (if you have access to a laptop) to write your own DMP
- What are the important areas for you?

Loughborough University

Knowing the Vocabulary – Data Management & Grant Capture

Gareth Cole, Loughborough University

Define the followings terms and phrases

- Data Protection Act
- Sensitive data
- Third Party Copyright
- Embargo period
- Creative Commons Licences
- Open/proprietary formats
- Data preservation
- Data sustainability
- Data access statement / statement on access to the underlying research materials
- Data Archiving / data publishing

Funder requirements

- Exercise on can you find your funder requirements and do you know what they mean!
  - AHRC
  - ESRC
  - H2020
  - Wellcome Trust
  - British Academy

Main UK funder requirements

- Data should be as open and as accessible as possible
- Legal, ethical, commercial exceptions allowed
  - E.g. breaching the privacy of your participants; breaching an agreement with a commercial partner; Patents
- Data should be preserved and sustained for a set (depends on funder…) number of years
- Data should be made available with a Persistent Unique Identifier (e.g. DOI) where possible

Publisher requirements

- Predominantly STEM publishers but others encourage e.g. T&F
- PLOS
- Royal Society
- Springer
- Nature Publishing Group

Funder pages with requirements

- AHRC
- ESRC
- Wellcome Trust
- British Academy
  - No one policy – sections in grant proposals
- H2020
- Leverhulme
  - “Proposals in which the balance between assembling a data bank or database and the related subsequent research is heavily inclined to the former” are not funded

Our outstanding campus

Largest single site green campus in the UK