Research cannot flourish if data are not preserved and made accessible. All concerned must act accordingly – Nature editorial
Aims of session

• To understand what can be considered as research data and why this is increasing in importance.
• To be aware of ethical and legal issues around research data
• To know how to create a simple data management plan
• To name files logically
• To understand the importance of backing up work
• To be able to find repositories to share their data and to choose suitable formats
• To understand what is meant by metadata
Research data?

https://padlet.com/s_glen/pa4xtz90f0os
Background

Big Science

Research replication crisis

Digital Humanities

Journals such as Nature expect data to back up articles

2011 RCUK statement on data principles

Royal Society promotes data sharing

EPSRC policy on data
Video clip – why think about data?

https://www.youtube.com/watch?v=N2zK3sAtr-4
What are the issues here?
Legal and ethical issues

• Data protection
• Consent – sharing
• Anonymising
• Freedom of information act
• Who owns the data
What to plan

DCC checklist:

1. Administrative data
2. Data Collection
3. Documentation and metadata
4. Ethics and Legal Compliance
5. Storage and Backup
6. Selection and preservation
7. Data Sharing
8. Responsibilities and Resources
Have a go

Look at the Bath data management plan then think about your own data under the same headings.
Research data lifecycle
Quality Control

- Instrument calibration
- Taking multiple measurements
- Following protocols in taking measurements
- Validation rules
- Using controlled vocabularies
- Expert validation
- Statistical tests to identify anomalous values
File naming

- Dates – best to do YYYYMMDD so records file in chronological order

- Don’t use spaces or characters such as &. Try CamelCase or join_with_underscores.

- Try to make names meaningful but reasonably short.

- If others will be using the files make sure they understand the system.

- Write instructions for yourself on how files will be named to keep consistency.
Backups

Our Disaster Recovery Plan Goes Something Like This...

HELP! HELP!
How to share data

https://www.re3data.org/  
https://creativecommons.org/choose/
Formats

- PDF
- Word
- CSV
- XML
Metadata

Metadata STICKS
Title: The 2015 Northern Ireland general election survey: measuring political change in a power-sharing polity
Persistent identifier: 10.5255/UKDA-SN-851957
Depositor: Jonathan Tonge, University of Liverpool
Principal investigator(s): Jonathan Tonge, University of Liverpool
Sponsor(s): Economic and Social Research Council
Grant number: ES/L007320/1

Subject categories
Politics
Regular opinion polls
Society and culture

Abstract

Coverage, universe, methodology

Dates of fieldwork: 05 January 2015 - 04 January 2016
Country: United Kingdom | (Northern Ireland only)
Geography: Northern Ireland (weighted sampling of all 18 parliamentary constituencies)
Spatial units: Wards (Electoral)
Westminster Parliamentary Constituencies
Observation units: Individuals
Kind of data: Numeric
Method of data collection: 1,810 completed post-General Election interviews, based upon completion of a questionnaire, were undertaken of a representative sample of the Northern Ireland electorate, using stratified random sampling, based upon postcode address files, of the 18 parliamentary constituencies in the region. Four attempts were made to contact each person randomly sampled before an interviewer found an alternative. The fieldwork was conducted by Social Market Research Belfast.

Administrative and access information
Date of release:
First edition: 09 October 2015
Latest edition: 09 October 2015 (minor amendments only)
Copyright: Jonathan Tonge, University of Liverpool
Availability: UK Data Service
Author/Rightsholder. (Year). Title of data set (Version number) [Description of form]. Retrieved from http://