DIY Research Data Management
Training Kit for Librarians

Data sharing

Anne Donnelly
Liaison Librarian
College of Medicine & Veterinary Medicine
College of Humanities & Social Science
anne.donnelly@ed.ac.uk
Data Sharing

What it is

“... the practice of making data used for scholarly research available to others.” [Wikipedia]

Who’s involved

- the data sharer
- the data repository
- the secondary data user
- support staff!
Drivers for sharing

Openness

Open science, open source, open standards, open knowledge, open government, open architecture, open content and ....
Drivers for sharing

Open Data!

“... a philosophy and practice requiring that certain data are freely available to everyone, without restrictions of copyright, patents or other mechanisms for control.”

[Wikipedia & elsewhere]
Drivers for data sharing

National research policies

• **Research Councils UK** “Publicly funded research data are a public good, produced in the public interest, which should be made openly available with as few restrictions as possible in a timely and responsible manner that does not harm intellectual property.”

• **Economic & Social Research Council** “… publicly funded research data … valuable, long-term resources that, where practical, must be made available for secondary, scientific research.”

• **Medical Research Council** “… publicly-funded research data … should be openly available to the maximum extent possible.”

• **Wellcome Trust** “… aim(s) to ensure that the data generated by the research we support is managed and shared in a way that maximises the benefit to the public.”
Drivers for data sharing

Institutional mission and policy

• **University of Edinburgh mission**  “The creation, dissemination and curation of knowledge.”

• **University of Edinburgh RDM policy**  “RDM plans must ensure that research data are available for access and re-use where appropriate and under appropriate safeguards.”
Drivers for sharing

Why not!

“... we have the technologies to permit worldwide availability and distributed process of scientific data, broadening collaboration and accelerating the pace and depth of discovery ...” [John Willbanks, VP Science, Creative Commons – in Wikipedia]
Barriers to data sharing

“Scientists would rather share their toothbrush than their data!” [Carole Goble, Keynote address, EGEE (Enabling Grids for EScienceE) ‘06 Conference.]

Barriers to sharing can relate to ...

- the Researcher - intellectual property issues
- the Institution - unrealised commercial value
- the Subject - confidentiality
Data sharing benefits ...

... Research & Society

- Avoids duplication of effort and resources.
- Makes publicly available that which has been publicly funded.
- Maximises return on research investment.
- Academic and scientific integrity
  - Increases transparency & accountability
  - Facilitates scrutiny of research findings
  - Prevents fraud
- Extends reach of original research.
- Facilitates and promotes new lines of enquiry.
- Fosters collaboration.
- Allows the creation of new datasets
- Provides datasets for education and training
Data sharing benefits ...

... the Researcher

• Facilitates compliance with funding council requirements
• Authenticates/validates research
• Enhances reputation through increased reach & impact
• Achieves increased visibility of research
• Assists with long term data storage
• Enables future retrieval of the data by the researcher and future researchers
Planning for data sharing

“Everyone in a research team should have a clear sense of their responsibilities in ensuring that ... research data are of the highest quality; ... are well documented so that other researchers can access, understand, use and add value to them ... independently of the original investigators.”

[MRC Guidance on Data Management Plans]

Issues to consider

- Future ‘shareability’ of the data: format, software, anonymisation, documentation, ethics, consent & confidentiality
- Description & documentation
- Future rights management & licensing
- Timescale for release
- Infrastructure for sharing
Ways of sharing data

Research-funder repository
- e.g. ESRC Data Store
  http://store.data-archive.ac.uk/store/

Institutional repository
- e.g. Edinburgh DataShare
  http://datashare.is.ed.ac.uk/

Discipline-specific repository
- e.g. Figshare
  http://figshare.com/
Thank you!