Personal, not Painful: Practical and Motivating Experiences in Data Sharing

Louise Corti
Collections Development and Producer Relations Team

Data Dialogues
University of Cambridge
28 July 2016
Overview today

- Enable safe access to data that meet the needs of data protection … yet fulfil the demands for open science and transparency
- Utilise appropriate policies, licencing and access pathways
- Examples of ‘difficult’ quantitative and qualitative data that have been successfully shared
- Helping empower data owners in their decision-making
Tensions in sharing one’s research data

✔ Want to do good science with integrity
✔ See the benefits of sharing and collaboration
✔ Need to meet funders policy/contractual mandate

✘ Sharing data ‘not possible’ - ethical, legal or practical reasons
✘ Unwillingness to share – my IPR, ‘they won’t understand’, no time

Let the counselling begin…
The counselling session

• Understand and don’t dismiss a data problem
• A data owner’s perspective is meaningful to them
• Our role is to listen, counsel, educate and persuade that something can be done
• It is always possible to share something useful from research
• Reel off the incentives
Collections Development work at UKDS

Trawling

Line-catching
The UK Data Service is lucky to build on almost **50 years of practice**

**7000 data** collections across social science spectrum

Supporting the ESRC Data Policy since 1995

Dealing with **1000s** of social science data creators using accumulated tried and tested **best practice**
Data policies

Protocols and templates

Legal agreements/concordats

Negotiation techniques

Incentives for sharing

Key contacts

Expertise and experts
Successful data appraisal

✓ Use Collections Development Policy – an evolving and living document

✓ Use Appraisal and Selection Criteria – agreed audit trail, research integrity, ethics and disclosure risk review

✓ Use an expert Data Appraisal Group with skills across the Service.
Access conditions

Depositor selects, with guidance, the access category most appropriate for the data

- **Open**
  - available for download/online access under open licence without any registration

- **Safeguarded**
  - available for download/online access to logged-in users who have registered and agreed to an End User Licence

- **Controlled**
  - available for remote or safe room access registered users whose research proposal has been approved by an access committee and who have received specialist training
Successful data ingest

✓ Standard depositors licences or data sharing agreements and End User Agreements – not bespoke
✓ Work to an agreed timetable for data publishing
✓ Use robust and explicit quality assessment techniques
✓ Use standard deposit metadata and data description
Not all data are sensitive

- Common worry about onward sharing of data covering ‘sensitive topics’
- But not ‘sensitive data’ for the DPA

- Deal with personal sensitive data properly
- Discuss consent status and implications
- Offer appropriate (multiple) pathways to access
  - Anonymise and/or control
- Last resort, offer permission to reuse/control access
Negotiation on access – **UK cohort studies**

- Safeguarded access: most versions

- Special Licence Access: some e.g. Millennium Cohort Study, 2001-2003: Hospital of Birth:— uncoded free-text hospital variable

- Secure access: some e.g. linked Department for Education educational records

  Approved Researchers; 5 Safes framework
Caveats of locking down

• Significant administrative burden:
  • 100s datasets – securing data access owners and committees
  • 1000s users – approved projects, approved researchers, bespoke output checking
• Future proof permission-granting owners moving on:
  • Advocate a Data Access Committee
  • Renegotiate access control level
Health and Social Consequences of the Foot and Mouth Disease Epidemic in North Cumbria, 2001-2003

SN 5407, Maggie Mort, Lancaster University

- Study funded by Department of Health
- Panel of 54 local people in affected area during FM crisis:
  - farmers, agricultural professionals, small businesses, health professionals, vets, residents
- Weekly diaries for 18 months
  - how their life was affected by the crisis
  - process of recovery observed around them (handwritten)
- In-depth interviews and group discussions (audio recordings, transcripts)
Data sharing steps

✓ At end of research – desire to archive data
✓ Met with UKDS for advice data archiving
✓ Sought advice from copyright specialist on archiving
✓ Developed separate consent forms for written and audio material, opt in / opt out and an embargo option
✓ Piloted discussion on data archiving with panel members:
  ✓ data anonymisation, confidentiality, copyright, ownership
  ✓ understanding of archiving and information required
  ✓ user options of archived data - scholarly / educational purposes
✓ Discussed archiving individually with each panel member
Data sharing outcome

- Safeguarded access: 40 interview and diary transcripts were made available for re-use by registered users
- 3 interviews & 5 diaries embargoed until 2015
- 7 panel members declined archiving their data
- Permission only: audio files archived
- **Exemplary user guide and data lists**
Managing suffering at the end of life: a study of continuous deep sedation until death

SN 850749, Jane Seymour, University of Nottingham, 2012

- 30 case studies using qualitative methods
- Interviews with staff and relatives closely involved in the care of deceaseds who received continuous sedation until death at home, in hospital or in hospices
- One of the few studies which seek to take a qualitative perspective on clinical decision-making in this area
- Research talking about delicate and painful issues
- Frameworks for ethical review
Data sharing steps and outcome

- ESRC grant so data sharing mandated
- Negotiation with UKDS
- Discussed anonymisation strategy
- Safeguarded access
- Anonymised transcripts shared
### AVAILABLE FILES

#### Data

<table>
<thead>
<tr>
<th>Accessible to:</th>
<th>Registered users only (safeguarded data)</th>
</tr>
</thead>
<tbody>
<tr>
<td>File or bundle content:</td>
<td>Data</td>
</tr>
<tr>
<td>File or bundle description:</td>
<td>Bundle of all data files (Community bases, hospice based and hospital based case studies)</td>
</tr>
<tr>
<td>File format:</td>
<td>application/zip</td>
</tr>
<tr>
<td>License:</td>
<td>UK Data Service End User Licence</td>
</tr>
<tr>
<td>File size:</td>
<td>11Mb</td>
</tr>
</tbody>
</table>

#### Documentation

<table>
<thead>
<tr>
<th>Accessible to:</th>
<th>Anyone (open access data)</th>
</tr>
</thead>
<tbody>
<tr>
<td>File or bundle content:</td>
<td>Documentation</td>
</tr>
<tr>
<td>File or bundle description:</td>
<td>The perspectives of clinical staff and bereaved informal care-givers on the use of continuous sedation until death for cancer patients: The study protocol of the UNBIASED study</td>
</tr>
<tr>
<td>File format:</td>
<td>application/pdf</td>
</tr>
<tr>
<td>License:</td>
<td>Creative Commons Attribution 4.0 International</td>
</tr>
<tr>
<td>File size:</td>
<td>187Kb</td>
</tr>
</tbody>
</table>
Handling queries on self-deposit

- 940 data collections published so far in ReShare
- Some hand-holding for depositors prior to upload
- Full-time repository administrator
  - Junior research level with MA in social science
- 1 in 10 questions relayed up to more senior staff
  - Ethics and disclosure review
  - Some formats and technical issues
- Query tracking system to manage/log/view responses
Sensitive case

- Famous academic anthropologist
- Unwilling to share ESRC data from years in the field
- Life’s works over 50 years
- On retirement…
I’m ready now to share that data you asked for…
Data sharing steps

- Ok to say ‘No’ but helpfully refer
- Take time to sift and weed if you see high value
- Discourage from such physical offers unless you can handle them!
- Be wary that saying no also might have consequences
Key messages from these cases

- Expertise in negotiating and listening
- Understanding of and confidence in discussing the research approach and data
- Appreciation that sensitive topics do not necessarily mean ‘sensitive data’ (DPA)
- Flexibility in presenting options to PIs, don’t offer the most locked down first
- Giving concrete advice and solutions – submit samples of data to assess
- Point to other successful cases
- Be ready to refer on!

LOVE research data
Make it easy to publish and upload data

Edit collection: Data Collection #851515

To deposit a data collection, you must accept the ReShare Terms and Conditions.

- I confirm that I am the owner of the copyright and associated intellectual property rights in the whole Data Collection or am otherwise lawfully entitled to grant this licence on behalf of each and every owner;
- I grant a non-exclusive, royalty-free licence to the UK Data Archive (a department of the University of Essex and not a separate legal entity) of Wivenhoe Park, Colchester, CO4 3SQ (the "University") to hold, make copies of, and disseminate copies of the Data collection, in accordance with the access conditions I will specify when uploading data files: open data accessible to users without registration; or safeguarded data accessible to users registered with the data services provided by the UK Data Archive.

I agree to the ReShare data deposit terms and conditions

Save for later  Cancel  Next >
Guidance and training resources

• Web-based guidance and FAQ - not read by all..
• Training and capacity building

ReShare help guidance

1. Prepare research data for ReShare
2. Deposit data in ReShare
   Register and log in
   Terms and conditions
   Grant information
   Responsible people
   Data collection descriptions
   Upload files
   Deposit collection
3. Further guidance

ReShare data review procedures

1. UK Data Service review procedures
2. External peer reviewer procedures

UK DATA SERVICE REVIEW PROCEDURES
The UK Data Service reviews each data collection submitted to the ReShare repository for disclosure risk, copyright breaches, validity of file formats and level of documentation before publishing the collection. The quality or scientific validity of research data is not evaluated.

All numerical data files and at least a 10% random sample of textual data files are reviewed.

Non-English language data files are checked if a staff member of the UK Data Service is familiar with the language.

Any concerns, changes needing to be made to data files, or requests for additional documentation are communicated to the data depositor for actioning.

Data review comments and actions taken are noted in the data collection metadata record, as provenance information for the data collection.

Review procedures
Data review and common issues

✓ Overall, a positive experience for most depositors
✓ Mostly good quality data and documentation

A few recurring issues after upload:

✗ Poor - or complete lack of – documentation/methods
✗ Limited descriptive metadata for the catalogue record
✗ Poor file names
✗ No reason for excluding files, for which fieldwork took place
Remedies

• Relay issues back to depositor - timely
• Accept nothing unless it comes with a clear ReadMe file that explains what the collection is about
• Add alerts within the system and common issues to guidance
• Incentives – promote high quality collections
<table>
<thead>
<tr>
<th>Collection period:</th>
<th>Date from:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 October 2008</td>
</tr>
<tr>
<td></td>
<td>Date to:</td>
</tr>
<tr>
<td></td>
<td>30 September 2012</td>
</tr>
<tr>
<td>Country:</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>Data collection method:</td>
<td>Collection of information from official documentation across several countries</td>
</tr>
<tr>
<td>Observation unit:</td>
<td>Organisations, Text units</td>
</tr>
<tr>
<td>Kind of data:</td>
<td>Alpha-numeric</td>
</tr>
<tr>
<td>Type of data:</td>
<td>Qualitative and mixed methods data</td>
</tr>
<tr>
<td>Resource language:</td>
<td>English</td>
</tr>
</tbody>
</table>

**Access and Administration**

**Data sourcing, processing and preparation:** Information from various sources, including the International Bureau of Fiscal Documentation

**Copyright holders:**

<table>
<thead>
<tr>
<th>Name</th>
<th>Email</th>
<th>Affiliation</th>
<th>ORCID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Devereux, Michael</td>
<td>Unspecified</td>
<td>University of Oxford</td>
<td>Unspecified</td>
</tr>
</tbody>
</table>

**Contact:**

<table>
<thead>
<tr>
<th>Name</th>
<th>Email</th>
<th>Affiliation</th>
<th>ORCID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Devereux, Michael</td>
<td><a href="mailto:michael.devereux@sbs.ox.ac.uk">michael.devereux@sbs.ox.ac.uk</a></td>
<td>University of Oxford</td>
<td>Unspecified</td>
</tr>
</tbody>
</table>

**Notes on access:** None

**Publisher:** Economic and Social Research Council

**Last modified:** 28 Apr 2014 22:42

**AVAILABLE FILES**

**Data**

- [CBT_Tax_database.xlsx](CBT_Tax_database.xlsx)

**Documentation**

- [CBT_Tax_database_description.pdf](CBT_Tax_database_description.pdf)
Life at the front end of a Data Service
UK Data Service deposit resources

- UKDS depositor case studies
- UKDS webpages and video on preparing data
- UKDS webpages on operating the ESRC Data Policy
- UKDS webpages, book and video on RDM issues
- UKDS ReShare guide/checking guidelines
- Depositing Shareable Survey Data brochure
- UKDS Collections Development Policy
- UKDS Selection and Appraisal Criteria
- UKDS Data Purchase Guidelines
- Call to action: Use of DDI metadata in survey production process
Keep connected with us

- Subscribe to UK Data Service list: 
  www.jiscmail.ac.uk/cgi-bin/webadmin?A0=UKDATASERVICE

- Follow UK Data Service on Twitter: @UKDataService

- Facebook

- Youtube: www.youtube.com/user/UKDATASERVICE
Contact

corti@essex.ac.uk
Collections Development and Producer Relations team
UK Data Service
University of Essex
ukdataservice.ac.uk/help/get-in-touch.aspx