Facilitate Open Science Training for European Research

RESEARCH DATA MANAGEMENT AND OPEN DATA

6-7 October 2015
University of Manchester, UK
HANDLING QUALITATIVE DATA - TRANSCRIPTION, LISTING (AND QUALIBANK)

Libby Bishop, UK Data Service
Overview for this session

• Transcription

• Data listing (documentation)

• Digitisation

• QualiBank
Data transcription

- translation between forms

- all transcription is:
  - representational
  - selective - can be multiple-perspective for video
  - interpretive
  - theoretical
Transcription template

Should:
• possess a unique identifier
• adopt a uniform layout throughout the research project
• make use of speaker tags - turn-taking
• carry line breaks
• be page numbered
• carry a document header giving brief details of the interview: date, place, interviewer name, interviewee details, etc.

Other considerations:
• cover page
• compatibility with import features of Qualitative Data Analysis Software (QDAS) e.g., Atlas.ti, MAXQDA, and others
Transcription and data sharing

- added issues to consider when transcribing for data sharing

- in what format will the transcript be accessed?
  - paper
  - digital file
  - web
  - standalone or part of collection

- who will be reading the transcript?
  - need for more/different contextual information (‘metadata’) for secondary users?
  - exposes the researcher’s practices
In practice: transcript format

<table>
<thead>
<tr>
<th>Study Name:</th>
<th>Interview number:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depositor:</td>
<td>Interview ID:</td>
</tr>
<tr>
<td>Interviewer:</td>
<td>Firstname Lastname</td>
</tr>
<tr>
<td></td>
<td>Date of interview:</td>
</tr>
</tbody>
</table>

**Information about interviewee**

<table>
<thead>
<tr>
<th>Date of birth:</th>
<th>Marital status:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender:</td>
<td>Occupation:</td>
</tr>
<tr>
<td>Geographic region:</td>
<td></td>
</tr>
</tbody>
</table>

Y=Interviewee

I=Interviewer

Y: I came here in late 1968.

I: You came here in late 1968? Many years already.

Y: 31 years already. 31 years already.

I: (laugh) It is really a long time. Why did you choose to come to England at that time?

Y: I met my husband and after we got married in Hong Kong, I applied to come to England.

I: You met your husband in Hong Kong?

Y: Yes.

I: He was working here [in England] already?
Data-level documentation

• Aim to embed this documentation in your data file

• Bigger issue lurking - metadata vs. data….no clear boundary (NSA)

• Qualitative data/text documents:
  • interview transcript speech demarcation (speaker tags)
  • document header with brief details of interview date, place, interviewer name, interviewee details, context
A user guide could contain a variety of documents that provide context: interview schedule, transcription notes, even photos.

**Interview structure**

- Your Photography
- Other People’s Photograph
- Your Photoblog
- Community
- About Your Site

**YOUR PHOTOGRAPHY**
Do you have any models who you look to? Favourite photographers (famous or otherwise)? Favourite styles of photography?

How would you say, in general, that you learn how to be a better photographer? Or is getting “better” not relevant to you? If not, what is important to you about your photography?

**OTHER PEOPLE’S PHOTOGRAPHS**
Where do you see other people’s photographs (whether “professional” or “amateur”)? Mostly online? Mostly offline?
Qualitative study - data listing

- Data listing provides an at-a-glance summary of interview sets

### Study Number 5407
Health and Social Consequences of the Foot and Mouth Disease Epidemic in North Cumbria, 2001
Mort, M.

The panel respondents for the study were divided into six population groups. The data list for the diary and interviews has been colour-coded accordingly for clarity, using the depositor's original colours:

- **Group 1: Farmers**
- **Group 2: Rural Business**
- **Group 3: Agricultural related occupations**
- **Group 4: Frontline Workers**
- **Group 5: Community**
- **Group 6: Animal / Human Health Professionals**

#### 1. Interviews

<table>
<thead>
<tr>
<th>Respondent ID</th>
<th>Population Group</th>
<th>Date of Birth</th>
<th>Gender</th>
<th>Occupation</th>
<th>Interview summary</th>
<th>Place of Interview</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM02</td>
<td>Group 6: Animal / Human Health Professionals</td>
<td>1975</td>
<td>M</td>
<td>Veterinary Surgeon</td>
<td>Family and background, career and work, arrangements during FMD epidemic and perceptions of situation</td>
<td>North Cumbria, respondent's home</td>
</tr>
<tr>
<td>PM03</td>
<td>Group 6: Animal / Human Health Professionals</td>
<td>1966</td>
<td>F</td>
<td>Veterinary Surgeon</td>
<td>Family and background, career and work, arrangements during FMD epidemic and perceptions of situation</td>
<td>North Cumbria</td>
</tr>
<tr>
<td>PM07</td>
<td>Group 6: Animal / Human Health Professionals</td>
<td>1964</td>
<td>F</td>
<td>Veterinary practice manager</td>
<td>Family and background, career and work, arrangements during FMD epidemic and perceptions of situation</td>
<td>North Cumbria, respondent's home</td>
</tr>
</tbody>
</table>
Digitisation of data

Non-digital data can (and should!) be digitised.

Approach dependent on situation
  - e.g. type of data, resources available, purpose of digitisation

Some general notes:
  Photographs
  - scan and save as TIFF
  Maps
  - scan, geo-reference using GIS software, and save as GeoTIFF
  Audio e.g. audio recording
  - capture as WAV
  Video
  - video formats complex, take care when digitising/converting to avoid degradation and errors
Digitising textual data

Text - more complex, with tiers of digitisation:

Create image file
- scan (or photograph) and save as TIFF image file
- used for poor typeface, handwritten materials, text with tables & graphs

Create searchable PDF
- collate TIFFs and convert to PDF
- bookmark PDF file for navigation: contents page, headings & metadata

Create rich text using Optical Character Recognition (OCR)
- automatically convert TIFF to RTF format
- requires rigorous proof reading and checking

Transcribe manually
- represent the original material as closely as possible
- avoid using formatting in data files
Exploring data online
**Interview with Ms. Ledbetter**  
SN5072 Mothers Alone, 1955-1966  

...bothered with little boxes for saving up. Food She pays quite a lot for meat and eggs she’s got potatoes from her brother-in-law who works on a farm but otherwise she had a lot out of the garden too. She’d not noticed any drop in the standard of living since her husband died he used to spend...
Interview with G2

FULL INTERVIEW TEXT

I: There’s just one or two factual questions first of all: do you mind me asking how old you are?

R: Well, I’m 51.

I: 51. And what schools did you go to yourself?

R: Well, I started off at Dunecot, and I finished at O’Neil Corse.

I: Whereabouts...

R: O’Neil Corse… well, it’s between Craigievar and Cunnie…

I: So you moved about…
Our data management guidance

- Online best practice guidance: ukdataservice.ac.uk/manage-data.aspx
- Managing and Sharing Research Data - a Guide to Good Practice: (Sage Publications Ltd)
- Helpdesk for queries: ukdataservice.ac.uk/help/get-in-touch.aspx
- Training: www.data-archive.ac.uk/create-manage/advice-training/events
Our guidance

- plan to share research data
- legal and ethical aspects of data sharing and reuse
- data copyright
- documentation and metadata to understand and use data
- file formats, organising, versioning and quality control
- storage, backup, encryption and security of data and files
- strategies for collaborative research
Other resources

• Research Data Management Training MANTRA (Edinburgh) - online learning units, also e.g. RDM guidance for N-Vivo
  http://datalib.edina.ac.uk/mantra/

• NASA Distributed Active Archive Centre:
  http://daac.ornl.gov/PI/pi_info.shtml

• Digital Curation Centre:
  • Data management planning
    http://www.dcc.ac.uk/resources/data-management-plans
  • Data management training / courses
    http://www.dcc.ac.uk/training/data-management-courses-and-training
Tools & templates

- Model consent form: http://www.data-archive.ac.uk/media/112638/ukdamodelconsent.pdf
- Survey consent statement: http://data-archive.ac.uk/media/147338/ukdasurveyconsent.doc
- Transcription template: http://data-archive.ac.uk/media/136055/ukdamodeltranscript.pdf
- Transcription instructions: http://data-archive.ac.uk/media/285633/ukda-example-transcription-instructions.pdf
- Transcription confidentiality agreement: http://data-archive.ac.uk/media/285636/ukda-transcriber-confidentiality-agreement.pdf
- Data list template: http://data-archive.ac.uk/media/2989/UK%20Data%20Archive%20Example%20Data%20List.pdf
- RDM costing tool: www.data-archive.ac.uk/media/247429/costingtool.pdf