
De-identification and pseudo-nymisation of qualitative data

Veerle Van den Eynden

UK Data Service

University of Essex

Managing and sharing research data: What is new with the GDPR?

4 May 2018

UK Data Service



Scenario

- Research on asylum seekers and refugees' experiences of forced labour, using interviews. These participants can be considered vulnerable.
- We want to use the collected research data (interviews) for analysis and share afterwards with other researchers
- How should we protect participant's anonymity?
- Read about this [case study](#)



Scenario - protect anonymity measures

- not record any official identifying data (e.g. Home Office numbers)
- let participants choose a pseudonym for interview
- password protect interviewee contact details
- do not connect in any way pseudonyms to the password protected interviewee contact details

Protect identities

- Avoid collecting identifying information
- Anonymise other information
- Plan in advance!



Anonymising qualitative data

- plan or apply editing at time of transcription
except: longitudinal studies - anonymise when data collection complete (linkages)
- avoid blanking out; use pseudonyms or replacements
- avoid over-anonymising – removing / aggregating information in text can distort data, make them unusable, unreliable or misleading
- consistency within research team and throughout project.
- identify replacements, e.g. with [brackets]
- keep anonymisation log of all replacements, aggregations or removals made – keep separate from anonymised data files

Anonymising qualitative data

Example: Anonymisation log interview transcripts		
Interview / Page	Original	Changed to
Int1		
p1	Spain	European
p1	E-print Ltd	Printing
p2	20 th June	June
p2	Amy	Moira
Int2		
p1	Francis	my friend

P31. Joan → Mary

P97. Carol → {Mother}

P34. Colchester → {Town in S.E.England}

P65. Welshpool High School → @@##High School##@@



What if anonymising is impossible?

- Obtain consent for sharing non-anonymised data
- Regulate or restrict user access



Exercise: qualitative data

Seymour, Jane (2010-2012). Managing suffering at the end of life: a study of continuous deep sedation until death. [Data Collection]. Colchester, Essex: Economic and Social Research Council. 10.5255/UKDA-SN-850749

