A critical realist theory of the project qualifications

Data management plan

Date and version of plan:Original version 05.06.2019; current version 28.06.2020

Researcher:Jed Stone

Project context:The research is being carried with UCL Institute of Education

**1. What data will you be working with?**

*File extensions*

* bib, docx, ds2, jpg, mp3, mp4, mx20, nvp, pdf, tex (and associated extensions), and xlsx

*Software*

* Canon scanner IJ Scan Utility
* Foxit Phantom PDF
* Jabref 5.0
* MAXQDA 20.0
* Microsoft Excel 2019
* Microsoft Outlook 2019
* Microsoft Word 2019
* MikTeX 7483
* NVivo 12.0
* Olympus Dictation Management System
* TeXstudio 2.12.22

*Hardware*

* Canon CanoScan LiDE 220 compact scanner
* Dell XPS 15 9550 computer
* Grundig Digta 7 PDM7020-12 digital dictaphone
* Integral Courier Dual FIPS 197 encrypted 64 GB USB 3.0 drive
* Integral Crypto FIPS 197 encrypted 64 GB USB 3.0 drive (x2)
* Olympus DS-9000 mobile dictation device

*Expected maximum number and size of files*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Type** | **Size** | **Number** | **Size x number** |
| Consent and assent forms  | pdf/jpg | 1 MB | 200 | 200 MB |
| Consent/assent spreadsheet | xlsx | 5 MB | 1 | 5 MB |
| Letters (to participants/parents) | docx | 0.1 MB | 200 | 20 MB |
| Interview notes | docx | 0.1 | 100 | 10 MB |
| Audio recordings (ds2, mp3) | ds2/mp3 | 25 MB | 100 | 2.5 GB |
| Interview guides | docx | 0.1 MB | 100 | 10 MB |
| Feedback form | pdf/jpg | 1 MB | 200 | 200 MB |
| Transcripts | docx | 1 MB | 100 | 100 MB |
| Interviewee transcript reviews | pdf | 1 MB | 100 | 100 MB |
| Data analysis files | nvp/mx20 | 100 MB | 2 | 200 MB |
| Case study documents | various | 1 GB | 1 | 1 GB |
|  |  |  |  | **4.5 GB (approx.)** |

*Estimated maximum storage volume needed*

* Available storage volume: 64 GB
* Available backup storage volume: 64 GB

2. How will you organise and describe your data?

Data and research records will be stored in a digital folder named {`data and research'}, comprising the following subfolders:



Audio files will be stored in sub-subfolders, and given the same name as the subfolder:



Transcripts will be named in order (T01, T02…) and include the pseudonym of each participant.



3. How will you store and backup your data during the project?

* Non-digital data and research records will be stored in an EN 1143-1 certified grade 6 safe on secure premises until securely digitized.
* Non-digital data and records will be digitised, as soon as possible after creation/receipt.
* Digital files will be stored on a 64 GB drive featuring AES 256 bit FIPS 140-2 hardware encryption (`drive 1’)
* A second 64 GB drive featuring AES 256 bit FIPS 140-2 hardware encryption (`drive 2') will be used for fortnightly data backup
* Data created on a dictation device featuring AES 256 bit encryption will transferred to drive 1 and then destroyed from the dictation device using
* The digital dictation device and drives 1 and 2 will be stored in the grade 6 safe when not in use
* All other data, record, analysis, and project files will be created on and saved to drive 1, and backed up on drive 2.
* Audio recordings only will be copied to a third drive featuring AES 256 bit FIPS 140-2 hardware encryption (`drive 3')

4. Legal, security, ethical and/or commercial questions to consider during your project

* The research complies rigorously with relevant legislation e.g. Human Rights Act 1998, Equalities Act 2010, DPA 2018, GDPR 2018.
* The research will take place in a secure setting: on premises with CCTV, alarmed buildings, and controlled entry.
* Arrangements for data security are outlined elsewhere in this document.
* Ethical questions are addressed in the separate ethics form.
* No commercial questions are raised by the project.
* Transcript data will be pseudonymised and strenuous effort will be made to further de-identify data using redaction or fictionalisation, whilst not compromising data integrity.
* Data could be sensitive, insofar as participants could disclose abuse or criminality.
* Sensitive data will be disclosed as appropriate in such instances. For example, data raising concerns about the welfare of any participant will be referred immediately to the designated safeguarding lead.
* Only transcribers will have access to the original data. Transcribers will need to sign a legally binding non-disclosure agreement approved by UCL data protection officer.
* Permission to use the data in reporting is requested in the consent and assent forms.

5. What data will you keep after the project? How long and where will they be archived?

* Data will be securely destroyed at the project end (estimated 25 June 2024).
* Digital data on all devices will be destroyed using the Gutmann 35 pass algorithm.
* Hard data will be destroyed using a micro cut shredder (DIN} security level 5).

6. How will others learn that your final datasets exist? How will they be able to access them?

* Datasets will have been destroyed and will not be available to others.
* Research outputs will be available on or via the website [www.questioningeducation.co.uk](http://www.questioningeducation.co.uk).
* The final full report will be accessible through UCL's electronic repository. It will thereby be in the public domain.