

Planning to archive your postgraduate research data: 6 things to bear in mind

by Irish Qualitative Data Archive, March 2012



The value of data archiving is increasingly recognised in Irish research policy. For example, the Irish Research Council (formerly IRCHSS) requires that "whenever data is to be collected with the support of a grant awarded by IRCHSS, applicants must specify the means by which that data will be made available as a public good for the use by other researchers". If you plan to deposit your qualitative research data within an archive, there are six aspects of best practice to bear in mind.

Please refer also to the document **Best Practice in Archiving Qualitative Data** which is available through the Irish Qualitative Data Archive at www.iqda.ie

1. Plan ahead

- Ensure that a comprehensive data management plan is in place at the onset of your project, and continue this through to the data archiving and dissemination phases of your project.
- Gain and record consent from participants to ensure that ethical and legal procedures are respected. Consent should include both consent to participate in the project and consent to archive the data. An IQDA sample consent form is available at <http://www.iqda.ie/sites/default/files/consentform.pdf>
- Ensure that data are collected in an open source formats such as Rich Text Format (.rtf) or Plain text data (.txt) for text files, and WAV (.wav) or Free Lossless Audio Codec (.flac) for audio files.
- Develop an anonymisation plan to preserve participant privacy and confidentiality. Identify texts that may be particularly sensitive and develop an archiving plan for such data. Common approaches include removal of sensitive text segments or placing a time embargo on dissemination of the archived document.

2. Use the available software

- Encryption software ensures that data are not accessible to unauthorised users. Common encryption programmes include TruCrypt, GnuPG and PGP.
- Automated anonymisation tools enable a standardised and rigorous anonymisation process for textual materials, but bear in mind that all documents must also be manually reviewed following the automated anonymisation process. The IQDA supplies an anonymisation tool at this location: [/sites/default/files/iqda_anonymisation_tool.jar](http://www.iqda.ie/sites/default/files/iqda_anonymisation_tool.jar)

3. Keep track of all changes

- This is a time-intensive process that is generally carried out by the researcher and it is wise to build in the costs associated with this process when applying for funding. This generally involves locating major identifying data (such as names) any identifying details (such as occupation details) within each text, and selecting replacement pseudonyms that reflect the significance of the original text. Create a 'Changes File' in which the pseudonyms that have been used to anonymise data are recorded.
- Use a clear and consistent unique identifier to indicate text that has been anonymised. IQDA guidelines on anonymising data are available at <http://www.iqda.ie/sites/default/files/AnonymisationProtocolV5.pdf>
- You may come across areas where there is doubt as to the best way to proceed with anonymisation of data. Keep track of any areas of such doubt in your 'Changes File' whereby such alerts can be reviewed at the end of the process, perhaps in consultation with a supervisor.

4. Make your data re-usable by future researchers

- During the process of anonymisation, sensitive statements that have been noted in the 'Changes File' will be altered using appropriate pseudonyms. If a statement cannot be anonymised the entire text may need to be removed and explicitly marked as such.
- Include additional contextual data (metadata) to facilitate re-use, by creating a header within each document that contains brief contextual details such as gender of those present, location of interview and so on.
- Prepare a 'User Guide' which gives an overview of the data to assist re-use in the future. An IQDA guide to material that may be included is available at: <http://www.iqda.ie/sites/default/files/Contextual%20Information.pdf>

5. Deposit the anonymised data with the archive

- Deposit the final anonymised document as a read-only file, which allows users to copy and paste as required but not to alter data.
- Create a dataset that contains all the sensitive segments that have been removed or changed. If possible deposit this dataset with a long-term embargo on access, so that in the future researchers will have access to the full data-set as the level of sensitivity reduces with the passage of time.

6. Disseminate the anonymised data via the archive

- Once deposited the archive should manage all user queries and access to the data on your behalf. The archive should have validation procedures in place for granting access to end-users of your data. An IQDA example of an end-user licence is available at http://www.iqda.ie/sites/default/files/IQDA_Data_Access_Request_Form.pdf
- The archive should have a standard citation protocol whereby end-user will acknowledge you as the data creator.