UK · DATA ARCHIVE

DATA MANAGEMENT FOR QUALITATIVE DATA USING NVIVO9

Contents

Scope

This document provides guidelines and suggestions for data management aspects that can be applied in NVivo 9 for qualitative research data. Suggestions both aim to ensure that research data are organised, documented and safeguarded well and that data can be shared with the wider academic community and remain valid in the long term. It also highlights which data management aspects may not be available or suitable in NVivo and therefore need to be considered separately. These guidelines are based on the UK Data Archive's current best practices in managing and sharing data.

The guidance applies to a wide variety of qualitative data, including textual data such as interview transcripts and field notes, audio, image and video data and their respective documentation.

1. Preparation before importing data into NVivo

1.1. Transcription and digitisation

Whilst data can be transcribed directly from audio into NVivo, the absence of spell-check facility in NVivo may negatively affect transcription quality. It is preferable for transcription to be carried out using word processing software and for text to be spell-checked before transcripts are imported into NVivo.

Care should be taken to use a uniform transcription format and style for all transcripts. Contextual information of an interview can be included in an initial paragraph. Researchers should avoid the use of headers or footers in transcripts, as those are stripped when importing data into NVivo. All information should be in the body of the text file.

Formatting and styling of transcripts (and contextual info) can be used for later auto coding in NVivo, e.g. using different heading styles for auto coding of questions, topic headings, speaker tags, etc.

Any non-digitally born data digitised to PDF before import, should be digitised using Optical Character Recognition (OCR) software such as Adobe Acrobat Professional or ABBYY FineReader, to be able to later annotate and code text.

1.2. Anonymisation of textual data

Whilst it is possible to edit and anonymise textual data in NVivo, it is preferable to carry out any initial anonymisation that can be done early in the research - such as replacing real names of people, organisations and locations with pseudonyms - before importing textual data files into NVIVO and before any coding is carried out. Word processing software enables the use of find and replace for possible editing of names throughout a transcript.

1.3. File naming

Before importing data into NVivo, consider the best way or structure to name data files, using standardised and uniform file names for each group of files. Project acronyms, events, dates, data types and participant IDs can be used to help identify files easily.

Whilst numerical-based file names provide good structure (e.g. Int001, Int002, RTF_Wave1_Interview01_01-12-2009), they may not be as user friendly as using name-based file names during analysis in NVivo (e.g. Int_Sam, Int_Marc).

2. Importing and organising data in NVivo

2.1. Organising files in folders

NVivo allows structured organising of imported files into source folders of Internals, Externals and Memos. The optimal way of organising data files into folders depends on the research and planned analysis. Folders can be created according to e.g. cases studied, locations, study waves, data types, etc. Internal folders can be created for various data files such as interviews, focus groups, field notes, audio recordings. One might wish to consider creating different folders for the original, anonymised and coded or annotated versions of data. At any time can sources can be re-structured without affecting coding. Documentation files can be organised in folders within Memos.

Where audio or video recordings take up too much space in the NVivo database, it may be preferable to organise them as Externals.

When data sources are exported from NVivo, e.g. for data archiving, the folder structure may provide important descriptive information about the data, as it shows the cohesion of and relationship between individual data sources. Whilst this structure of sources cannot be exported as such, a 'print screen' of the structure of folders can be saved as an image or PDF file. Information about the structure of folders can be imprinted as additional information within an exported file when exporting Internals and Memos, if the appropriate options are selected during export.

Folder structure info can also be included as attributes of sources.

2.2. Data import

Textual data files such as interview transcripts and field notes, as well as textual documentation files, can be easily imported into (and exported out of) the NVivo database, in MS Word (.docx, .doc), Adobe PDF (.pdf), Rich text (.rtf) or plain text (.txt) format. Imported sources become editable text files in NVivo.

How NVivo9 imports PDF files and whether or not they are editable after import depends on whether they were born digitally - PDF created from an original word document or webpage - or not - PDF created from scanned images or documents. If scanned text is digitised using Optical Character Recognition (OCR) it will be partially editable; otherwise it is not..

Tabular data can be imported in MS Excel (.xlsx, .xls) and comma- or tab-delimited text files (.txt) format. Also MS Access or SQL Server tables can be imported in these formats.

Audio files can be imported in MPEG-1 Audio Layer 3 (.mp3) or Microsoft Windows Media (.wma, .wav) format.

Video files can be imported in MPEG file formats (.mpg, .mpeg, .mpe, .mp4), Windows Media (.avi, .wmv) or QuickTime (.mov, .qt) format.

Images can be imported in Windows bitmap (.bmp), Graphic Interchange Format (.gif), Joint Photographics Expert Group (.jpg, .jpeg) or Tagged Image File (.tif, .tiff) format.

One may want to consider whether or not audio, video or image files will later need to be exported or not, e.g. when archiving the entire data collection. Audio and video files can only be exported in the same format as they were imported in. Imported image files can only be exported in .jpg format, irrespective of the import format. It may be better to keep separate copies of data files for archiving outside the NVivo database.

From a point of view of long-term preservation, recommended formats for import are .wav for audio files and .mp4 for video files.

Researchers may want to consider whether it is best to import audio, video or images as Internals or whether to keep them as Externals, due to the size of the resulting NVivo project file if numerous large files are imported. Externals are best kept in a single folder on the PC or server.

3. Editing data after import

3.1. Anonymisation and editing

Whilst not ideal, it is possible to carry out anonymisation and editing of textual data in NVivo using the Formatting tools on the Home ribbon.

Basic find and replace edits can be carried out within individual sources, not across all sources. To search for terms to anonymise across sources, text search queries can be used. Whilst one cannot directly replace text from a text search query, one can use such query and create the results as a new node, then edit individual sources using find and replace within each source.

Alternatively, during coding and analysis of data, any text that may require later anonymisation can be flagged up and coded, creating an 'anonymisation' node to hold such anonymisation coding. This can then easily be checked when anonymisation is carried out. The anonymisation node can be removed when it is no longer needed.

Note that there is no spell-check facility for textual information.

3.2. Version control

In server-based NVivo, 'Edit mode' ensures that team members do not make conflicting changes to source content. A source must be open in edit mode to be able to modify the content—for example, edit the text in a document source, or add transcript rows to a video. When a source is open in edit mode, the user has exclusive control of the content. Other users can view, code, annotate, link or query the source, but cannot change the content.

4.   Data description and documentation

4.1.   Documentation

Documentation files like the methodology description, project plan, interview guidelines and consent form templates can be imported into the NVivo project file and stored in a 'documentation' folder in the Memos folder or can be linked from NVivo9 externally. Additional documentation about analyses or data manipulations can be created in NVivo as memos.

A date and time stamped project event log can record all project events carried out during the NVivo project cycle, with date stamps and user logs. This log provides important information about the data sources, editing done, coding and analysis carried out, etc. A log may loose its records when different versions of a project file are merged. The log file can be exported as a MS Excel, PDF or RTF file.

Additional descriptions can be added to all objects created in or imported to the project file, such as the project file itself, data, documents, memos, nodes and classifications.

All textual documentation compiled during the NVivo project cycle can later be exported as textual files. Classifications and event logs can be exported as spreadsheets to document preserved data collections. The structure of the project objects can be exported in groups or individually. Summary information about the project as a whole or groups of objects can be exported via project summary extract reports as a text files, spreadsheet or xml file.

To facilitate future data archiving, it may be advisable to keep all essential documentation of an NVivo project in a 'documentation' folder, for easy export when data are archived.

4.2.   Classifications and attributes

Researchers can create classifications for persons (e.g. interviewees), for data sources (e.g. interviews) and for coding. Classifications contain attributes such as the demographic characteristics of interviewees, pseudonyms used, and the date, time and place of interview. A source classification records information about the type of source in a project, with some predefined classifications available in NVivo for interviews and focus groups. A node classification records demographic or descriptive information about nodes that are either persons, organisations, communities or other cases in a project.

Classifications can be manually created within NVivo, or imported from an external classification sheet (.xls) or from another NVivo project file. If generic classifications are created beforehand, attributes can be standardised across all sources or persons throughout the project. Attributes can be added or modified at any time.

It is possible to list pseudonyms and corresponding real names amongst the attributes of a classification for people or organisations, which may help during analysis.

Classifications can be used to create a data listing that provides an overview of all data sources. Data listings are used by the UK Data Archive to accompany a collection of interview transcripts.

4.3.   Consent for data sharing

For data archiving it is important to know whether or not consent for data sharing has been obtained from a participant. This can be recorded amongst attributes for people interviewed. Template consent form(s) and information sheet(s) used in the research can be incorporated as documentation files in the NVivo project file, alongside other documentation.

Researchers should not upload individual signed consent forms as they contain personal data and therefore need to be stored securely.

5.   Storage, back-up and access

5.1. Data storage

All internal sources, memos and objects that are created in NVivo 9 during analysis, including nodes, classifications, annotations, links, relationships and sets, are stored in a single NVivo 9 project file with an .nvp extension. There is virtually no limit to the objects or sources you can create in NVivo 9. This single file should be stored securely on a PC or a server. NVivo is by default set up to prompt the user every 15 minutes to save the project file; something which the user can change according to needs in File>Options>Notifications.

Researchers may want to store master versions of all data files such as interview transcripts on a PC or server, as a back-up.

External sources in an NVivo file are not stored in the NVivo file itself and need to be stored safely elsewhere. It may be good practice to organise all externals for an NVivo project in a single folder on the PC or server. This makes storage and back-up easier.

5.2. Back-ups

Researchers need to make sure to carry out regular back-ups of the NVivo file. If the file is held on an institutional server, the institutional back-up strategy may take care of that. A separate strategy may be needed for externals.

5.3. Access controls

Both single-user and server-based NVivo 9 project files can be password-protected or read/write permissions can be set. The project owner can control who can view or modify the project, by allocating reader or contributor status to users.

No differing levels of access conditions or password protection can be set for individual folders, objects or sources within a project file. This means that all users with contributor status have access to all content of the NVivo file. The project event log, however, can track events such as modifications by individual users. For collaborative projects consideration may need to be given whether that is desirable or whether separate project files may need to be created.

Data that need to be handled securely such as personal data and interviewee's real names may need to be excluded from the NVivo project file in some collaborative server-based projects.

6.   Exporting data and documentation files

Exporting data and documentation may be important when data are being archived for further re-use. Whilst master versions of data files may have been held elsewhere, it is very likely that useful documentation for the data will need to be exported from the NVivo project file. It is preferable for long-term preservation to archive individual data and documentation files, rather than the entire NVivo project file.

6.1. Exporting data

Textual data can be exported from an NVivo project as MS Word (.doc, .docx), RTF, .txt or PDF files. In preparation for archiving, sources are best exported as RTF files. Files can be exported as individual files. It is also possible to export a group of sources such as interviews, whereby each source will be saved as an individual file within a specified folder. Information about the structure of folders can be imprinted as additional information within an exported file when exporting Internals and Memos.

Tabular data such as incorporated surveys can be exported as MS Excel or tab delimited text files.

Images can only be exported as .jpg files, which is a limitation. The UK Data Archive recommended preferred format for archiving image files is .tiff.

Audio files can be exported in various formats: .mp3, .wma or .wav. An audio file can only be exported in the same format as it was imported in.

Video can be exported in .mpg, .mpeg, .mpe, .mp4, .av, .mwv, .mov or .qt format; again in the same format as the import format.

6.2. Exporting documentation

All textual documentation compiled during the NVivo project cycle such as memos, log files, documentation files in Internals can be exported as RTF textual file, or in MS Word, PDF or .text.

Classifications and project event logs can be exported as spreadsheets or tab delimited text files to document data collections. A group of classifications can be exported simultaneously, whereby they are exported as a single Excel file containing multiple worksheets. This is done by highlighting all, then Right-click>Export>Export classification sheets.

The structure of the project objects can be exported in groups or individually. Summary information about the project as a whole or groups of objects can be exported via project summary extract reports as a text files, spreadsheet or xml file.

Metadata such as source descriptions, annotations, relationships, links and memo links can all be exported. This may be important documentation for data archiving.

6.3. Preferred export formats

UK Data Archive preferred formats for data and documentation export for archiving	
Textual data	.rtf, .txt, .html
Tabular data	.xls, .txt
Image	.jpg (not ideal – keep preservation copies of images separate from NVivo in. tiff format, if available)
Audio	.wav, .mp3
Video	.mp4
Textual documentation	.rtf, .txt, .pdf, .html
Tabular documentation	.xls, .txt

7.   SUMMARY - best practice recommendations
- Transcribe interviews in word processing software before import to NVivo
    - apply uniform transcription format
    - do not use headers or footers
- Basic anonymisation in word processing software before import to NVivo
- Uniform, structured file names
- Plan how best to organise data files (sources) in folders within Internals
- Preferred data import formats, in view of later export for archiving
    - .doc or .rtf for text
    - .wav for audio
    - .mp4 for video
    - .tiff for images
- Decide whether to include images, audio or video as internals or externals
- Organise all External files in a single folder on PC or server
- Organise essential documentation files of the research in 'documentation' folder in Memos
- Use memos to document research
- Use classifications to document and annotate sources, persons interviewed, study cases, etc.
- Enable time-stamped project event log
- Record consent for data sharing amongst person attributes
- Store NVivo project file and external files securely
- Arrange regular back-up of NVivo file and externals

- Arrange read/write permissions for the project file
- Consider that all users have access to all content of an NVivo project file
- Consider best export formats for data and documentation in view of data archiving:
  - .rtf, .txt or .html for textual data and documentation
  - .xls or .txt for tabular data and documentation
  - .wav or .mp3 for audio
  - .mp4 for video
- Export all relevant documentation to accompany data files