



MANTRA NVIVO 9 DATA HANDLING EXERCISE

Introduction

Authors : Dr. Kakia Chatsiou (KC)
Contributors: Dr. Libby Bishop (LB), Louise Corti (LC), Dr. Veerle Van Den Eynden (VV),
Laurence Horton (LH), Robin Rice (RR), Dr. Bethany Morgan-Brett (BMB),
Dr. Kate Orton-Johnson (KOJ)
Publisher : EDINA and Data Library, University of Edinburgh
Project: Research Data MANTRA, www.ed.ac.uk/is/data-library-projects/mantra
Funder: JISC Managing Research Data Programme (2010-11)
Version: 00.13-Intro-00.04
Last updated: 30/08/2011

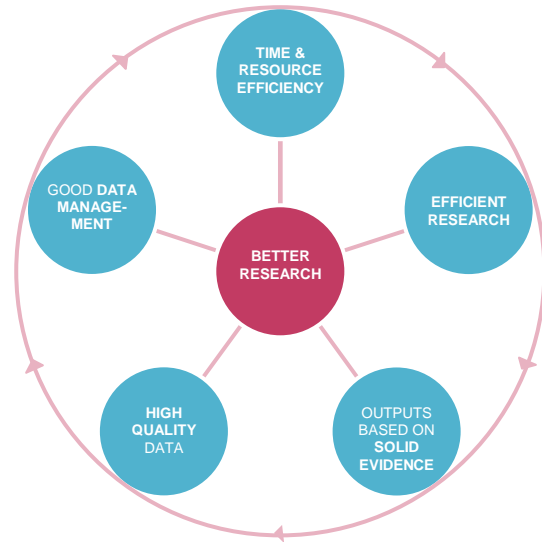
Contents

1	INTRODUCTION: SHARING AND MANAGING DATA WITH NVIVO 9	3
	0-1.....	4
	EXERCISE LEARNING OBJECTIVES.....	4
	EXERCISE TARGET AUDIENCE.....	4
	EXERCISE PREREQUISITES.....	4
2	ABOUT THE EDWARDIANS DATA COLLECTION USED IN THIS EXERCISE.....	5
	NEXT ► UNIT 1: PREPARING YOUR QUALITATIVE DATA FOR USE IN NVIVO 9.....	6
	NVIVO 9 EXERCISE DEMO FILES LISTING.....	7
	DATA MANAGEMENT WITH NVIVO 9 CHECKLIST	8

1 Introduction: Sharing and managing data with NVivo 9

Good data management is the foundation for good research. If data are properly organised, preserved and well documented, and their accuracy, validity and integrity is controlled at all times, the result is high quality data, efficient research, outputs based on solid evidence and the saving of time and resources. Researchers benefit greatly from properly managing their research data; that's why data management should be planned from the start of research. If it becomes part of standard research practice, then it should not require much additional time or incur further costs.

Data management primarily occurs within the lifecycle of a research project and should ideally be carried out by all members of the research team. When it comes to sharing research data, good management is essential to ensure that data can be preserved and remain accessible in the long-term, so they can be re-used and understood by other researchers. When managed and preserved properly, research data can be successfully used for future scientific and educational purposes, thus maximising the investment made in generating the data and increasing the visibility of the research.



For more information about good practices in data managing and sharing, see the *UK Data Archive Managing and Sharing Data Guide* (<http://www.data-archive.ac.uk/media/2894/managingsharing.pdf>)



It is recognised that different types of data created and managed across the research discipline spectrum may require certain discipline-specific approaches to data managing and sharing, and that projects may differ in their approach to specific data management and preservation issues.

This exercise has been written for researchers spanning the natural and social sciences and humanities who are using NVivo 9 as their main Computer Assisted Qualitative Data Analysis Software (CAQDAS) software. It was created as a **self-study exercise** that gives an overview of using NVivo 9 for managing qualitative data primarily with a focus on the pre- and post-analysis stages and developed as part of the MANTRA project exercises, available from <http://data.lib.edina.ac.uk/mantra>.

The exercise has also been designed to complement the University of Edinburgh's School of Social and Political Science postgraduate course 'Core Qualitative Data Analysis for Social Research', or other similar courses in research design, qualitative data analysis and data collection. It is not a self-contained introduction to NVivo 9 – users not familiar with the software might wish to refer to the NVivo 9 support pages (http://www.qsrinternational.com/support_getting-started.aspx) for more information.

The exercise consists of three units: Unit 1 guides you through **preparing your qualitative data** for use in NVivo 9 at the end of your project; Unit 2 provides guidance on **data management issues when working in NVivo** and Unit 3 guides you through the steps of **managing your data during post-analysis**.


In particular, **Unit 1** covers issues such as informed consent, Intellectual Property Rights (IPR) and copyright, good practices in formatting and anonymising transcripts and

recommended practices in structuring your data before you import them to the software. **Unit 2** includes advice on which data to import in your NVivo 9 project file and which are better left outside, guidance on filenames and file formats, logging events and keeping notes, tips on structuring your nodes and classifications and on data storage, data backup and access control. **Unit 3** guides you through recommended practices in exporting data, coding and information about your data including recommendations for file formats and outputs for sharing and archiving and suggested practices in managing and exporting your bibliographical resources in NVivo 9.

The material is complemented by a *Data Management with NVivo 9 checklist* for reference (at the end of the introduction and as a memo in the NVivo 9 project file included with the files) as well as a set of activities included in each section. Finally, in the **Memos>ExerciseAppendices** section of the NVivo9 project file included with the exercise files (2000-nvivo9-demo.nvp), we have included notes on some issues relevant to researchers such as **team working** and **setting up automatic backups**.

Throughout this document, the following conventions are followed:

- key notions throughout the document are marked with bold font e.g. **Data Management**
- filenames and code are marked with typewriter font e.g. 2000int001.rtf
- a sequence of clicks, steps or keystrokes to be followed are marked with a bold typewriter font e.g. **File > Save As..., Ctrl+Q**
- Tasks are in blue shaded border like this:


0-1.

Do this.

Exercise learning objectives

By the end of this exercise you should be familiar with:

- good data management practices and how to apply them while working in NVivo 9
- good practices in importing and exporting your data using NVivo 9
- how best to organise the structure of your sources, nodes and classifications in NVivo 9
- keeping a project journal and documenting your research using NVivo 9
- the limitations of NVivo 9 and some workarounds.

Exercise target audience

This exercise is aimed at:

- postgraduate students with some familiarity with NVivo 9 and sound knowledge of research design, qualitative data analysis and data collection who would like to know more about using NVivo 9 for data management
- research professionals with similar interests

Exercise prerequisites

- to have attended a course in research design, qualitative data analysis and data collection
- to have some familiarity with the NVivo 9 software (such as opening and accessing your project file and importing/exporting data). This guide is not an introduction to NVivo 9.

2 About the Edwardians data collection used in this exercise

The Edwardians data collection (Study Number 2000 from the Economic and Social Data Service (ESDS)) is a qualitative study, comprising of 453 qualitative interview transcripts, available as fully readable, searchable .rtf files. A quantitative data file was also developed during a follow-up project to code some of the data.

The major part of the collection consists of life-story interviews originally collected as part of the study *The Edwardians: Family Life and Work Experience Before 1918*. The interviews were undertaken in the late 1960s and early 1970s and formed the basis of the first national oral history project in the United Kingdom. A total of 453 interviews were recorded on reel-to-reel audio tape and later transcribed as typed, paper documents. The interviews were open-ended (guided by a schedule) and were between one and six hours long. In the latest version of the collection all interviews are accessible as .rtf documents and a set of .pdf cover sheets containing a two page summary of each interview is available.

The interview schedule covered topics such as: domestic routine, including the roles of husbands and children; meals; the upbringing of children; emotional relationships and values in the family; leisure; religion; politics; school; courtship and marriage; the wider family; relationships with neighbours and perception of the community structure; experience of work and occupational history of the whole family.

A related project, *Systematic Analysis of Life Histories*, is also included in the full dataset, which is available for download from the ESDS (<http://www.esds.ac.uk/findingdata/snDescription.asp?sn=2000>). The aim of the project was to prepare interviews collected for the *Family Life and Work Experience Before 1918* study for numerical coding shortly after completion. Additionally, this quantitative component of the collection has been enhanced. A selection of the original variables was cross-referenced for verification. The data set was then expanded to include a set of additional variables on occupation and location as well as textual summaries of the interview transcripts.

The original study materials were initially archived, catalogued and disseminated by ESDS Qualidata. Full paper/microfiche transcripts are held at the National Social Policy and Social Change Archive at the University of Essex and the original sound recordings are deposited at the British Library National Sound Archive and are available as .mp3s from ESDS. PDF & XML copies of the interview transcripts are available from ESDS Qualidata on request.

Further information on the project is available from the ESDS Qualidata Online: The Edwardians web page (<http://www.esds.ac.uk/qualidata/online/data/edwardians/introduction.asp>). More information, including how to access the full data collection is available from ESDS (<http://www.esds.ac.uk/findingData/snDescription.asp?sn=2000>).

For the purposes of this exercise, we will use a subset of this dataset including the following transcripts: 2000int001, 2000int002, 2000int003, 2000int016, a short version of the data listing relevant to these interviews (2000ulist.xls), as well as the study user guide (under 2000uguide.pdf in the full dataset download) split in many files. Since the documentation for this study was in .pdf format, we have included some .rtf versions too, in order to demonstrate how to use memos for documentation. The subset of the dataset for this exercise can be downloaded from the course website and unzipped: 2000-NVIVO9-demo_raw-data.zip. Note that one of the files within is itself a zip file: 2000-NVIVO9-demo.zip.

This data collection should be cited as:

Thompson, P. and Lummis, T., *Family Life and Work Experience Before 1918*, 1870-1973 [computer file]. 7th Edition. Colchester, Essex: UK Data Archive [distributor], May 2009. SN: 2000.

Next » UNIT 1: Preparing your qualitative data for use in NVivo 9

The next unit deals with issues you need to consider when preparing your data to work with NVivo 9, covering the following:

- informed consent
- Intellectual Property Rights (IPR) and copyright issues
- suggested practices in transcription
- suggested practices in anonymising your data
- recommended structure of folders for your data before import to NVivo 9

Available from MANTRA-NVIVO-exercise-Unit1.doc

NVivo 9 exercise demo files listing

Here is the tree folder structure of all the files that are used in this exercise:

```

.
| 2000int016.doc
| UKDA_Study_2000_information.htm
| 2000-NVIVO9-demo.zip
| dirlisting.txt
| 2000datasetsubset{1,2,3,16}.sav
|
\---2000-NVIVO9-demo_raw-data
    | 2000ulist.xls
    | 2000-nvivo9-demo.nvp
    | 2000bibliography.txt
    |
    +---doc
        | Edwardians - Interview schedule.URL
        | Edwardians - Themes.URL
        | Edwardians - Interview notes.URL
        | 2000dataset_data-dictionary.rtf
        | Edwardians - User Guide - 2000guide.pdf.URL
        | Edwardians - Occupational and geographical classifications.URL
        | 0.Follow_the_links_to_download_the_documentation_files.txt
        |
        +---video
            | 0.Follow_the_links_to_download_the_video_files.txt
            |
            +---transcript_formatted
                | 2000int001.rtf
                | 2000int002.rtf
                | 2000int003.rtf
                | 2000int016.rtf
                |
                +---transcript_originals
                    | 2000int001.rtf
                    | 2000int002.rtf
                    | 2000int003.rtf
                    | 2000int016.rtf
                    |
                    +---pictures
                        | 2000img01-brown_brothers_january_1910.jpg
                        | 2000img02-brown_brothers_fittings_department_600dpi.jpg
                        | 2000img03_davies_family_1911.jpg
                        | 2000img04-dr_frederick_brown_director_s_office.jpg
                        | 2000img05-fred_mills_1920.jpg
                        | 2000img06-leisure002.jpg
                        |
                        \---audio
                            | 2000int178.mp3.URL
                            | ESDS Qualidata - List of Edwardians audio clips.URL
                            | 2000int120.mp3.URL
                            | 0.Follow_the_links_to_download_the_audio_files.txt
                            |

```

Data Management with NVivo 9 Checklist

General and NVivo 9-particular Data Management Checklist	Notes
Consent	
Will you be collecting personal, confidential or sensitive information?	
Are you gaining informed consent for participation from research participants?	
Are you discussing data sharing or archiving with the informants from whom you collect the data?	
Are you gaining consent from informants to share data beyond your research?	
Are you asking for written or verbal consent?	
Have you explained the data sharing benefits to participants?	
Have you avoided uploading signed consent forms as internals to NVivo 9 project file?	
Have you recorded consent status as an additional attribute in the classification sheet?	
Anonymisation	
Do you need to anonymise data e.g. to remove identifying information or personal data? You will need to think about when this will be done and how.	
Have you removed all sensitive material and replaced them with pseudonyms before import to NVivo9?	
Have you created a backup copy of your originals in a secure storage space before anonymising them?	
Copyright	
If you are purchasing or re-using someone else's data sources have you considered how that data will be shareable?	
Have you established who owns the copyright of your data? Might there be joint copyright?	
Before archiving your data, have you ensured that all sources to be distributed to third parties have cleared copyright?	
Documentation	
How will you label and organise your data and files?	
Are your data self-explanatory in terms of variable names, codes and abbreviations used? Would someone else be able to understand your data?	
Do you have documentation which explains what your data means or gives context to your data? i.e. where an interview was carried out and with whom? variable lists?	
Have you considered how you will ensure the quality of your data? i.e. Data cleaning, checking transcriptions, data entry validation etc.	
Formatting and version control	
Which data formats will you use? i.e. .rtf, .xls Are these data formats likely to still be available to use in 10 years time?	
If data are held in various places, how will you keep track of versions?	
Will you know what the master version of your data files is?	
Transcription	
What conventions will you use for transcribing the data?	
Who will do the transcription?	

General and NVivo 9-particular Data Management Checklist	Notes
How will you ensure the confidentiality and security of your data during the transcription process?	
Have you spell-checked you transcripts?	
Security	
If your data is held on a USB stick or other mobile storage device how will you ensure the safety and security of the data on it?	
Are your files backed up regularly and are back-ups stored safely?	
How will you securely store personal or sensitive data? (both digital and hard copy data)	
Do you need to keep secure all or just part of your data?	
Do you have plans for how your data will be preserved once you have finished your research?	
Data access	
Who has access to which data during and after research?	
Are various access regulations needed?	

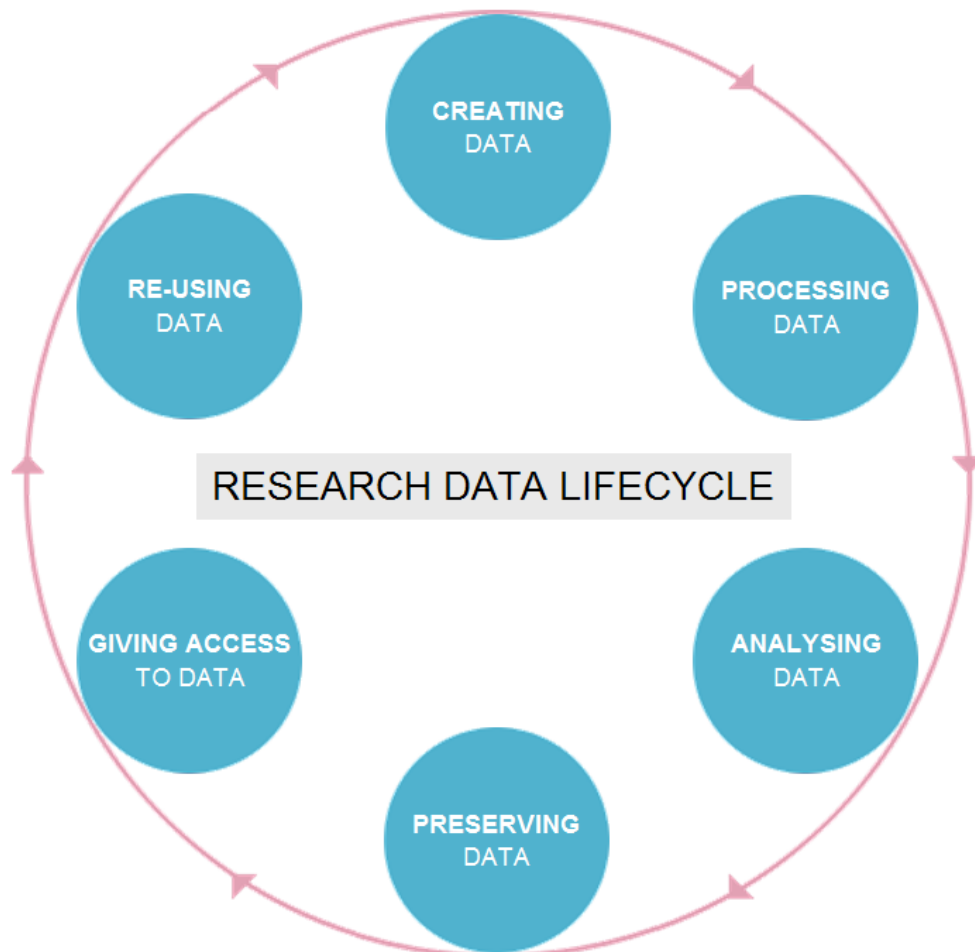


Figure 0.1 Research Data Lifecycle