DATA CURATION NETWORK
NVivo Data Curation Primer

Author: Hannah Hadley - hhadley@psu.edu
DCN Mentor: Dave Fearon - dfearon@jhu.edu
Affiliate Contributor: Margarita Corral (Atlas.ti primer) - mcorral@brandeis.edu

Topic	Description
File Extension	.nvp (Windows) .nvpx (Mac) .nvcx (NCapture tool: a web browser extension)
MIME Type	
Structure	
Versions	NVivo 12 for Mac NVivo 12 for PC (Pro, Plus) See appendix for extended table of previous versions.
Primary Fields or Areas of Use	● Qualitative and mixed methods research design ● Social Sciences, Humanities
Source and Affiliation	QSR International
Metadata Standards	● QuDex ● DDI
Key Questions for Curation Review	● Where will this data be housed? (e.g. enclave or other) ● Does the dataset contain components that cannot be ethically shared with the public? (e.g. human subjects; copyrighted materials or similar.) ● Is additional documentation needed to share components that are necessary for reuse? ● What components of the project (exported from NVivo) are ideal for reuse and preservation?
Created by	Hannah Hadley - hhadley@psu.edu
Date Created	April 2019
Date Updated and Summary of Changes	

Suggested Citation: Hadley, Hannah. (2019) NVivo Data Curation Primer. http://hdl.handle.net/11299/216583

Table of Contents

ReadMe: The scope of this primer

This primer is intended to provide an overview of the NVivo format to assist data curation actions. NVivo is a proprietary format of qualitative data analysis software (QDAS). **NVivo projects can be expected to contain sensitive data within their various components. While this primer does include some mention of de-identification, it is not comprehensive and does not describe all considerations of sensitive information, copyright or permissions.** Additional data curation primers are expected to be published on data sensitivity topics, and the author will link to any relevant sources once these become available.

Supplementary NVivo primers

Supplementary NVivo primers are expected to include more depth and detail for curating this format. Example topics of what a supplementary NVivo primer may include:
- Considerations for curating specific qualitative or mixed-method design in an NVivo project.
- Focus on curating specific areas or purposes of the format.
- Case studies or researcher interviews that inform our curation practices for NVivo, or that provide recommendations based on real-world examples.
- Detailed or step-by-step description of metadata for NVivo.
- Special considerations for curating older NVivo versions, or curating these from an archival (preservation) perspective to be reusable for current research.

Supplementary primers will be linked to this primer in the appendix section and linked to any appropriate subject section.

Description of the format

Overview

Nvivo is a proprietary format of qualitative data analysis software (QDAS) that can assist analysis of qualitative data for use with qualitative or mixed-methods research. Nvivo is used to organize, manage and visualize qualitative data, and may assist in generating queries and reports of data captured to its containers. Data may be captured by importing it directly or using the NCapture tool, it may also be entered manually as memos, journals, or nodes (which are coded references). NVivo files are a whole project with many independent and interrelated parts; not just a dataset. Nvivo may be conceptualized as having many containers to store a variety of data components.

NVivo components

An Nvivo project has multiple containers that may hold many different data types or objects including:
- **Nodes**: coded references to qualitative data components; classifications, attributes, values. (Note: Codebooks may be exported from nodes in NVivo versions 11 or later.)
- **Models**: graphic visualizations of information using NVivo modeller.
- **Sources**: All of the information that may be analyzed using this software. (Note: Sources can be in many forms; such as: text, tabular, audio, video, etc.)
  - **Internals**: various sources that comprise primary information for a project.
    - Annotations on text: memos within sources.
  - **Memos**: reflections about a project or its components that provide contextual information about a project.
    - Journals: whole project memos.
    - Memos (notes) linked to a source.
    - Memos (notes) linked to a node.
    - Internal linkages: links to sources or nodes from within sources or nodes.

- ○ **Externals**: proxy documents that could not be imported into a project.
  - ■ Hyperlinks: links to sources that are external to the project.
  - ○ **Framework Matrices:** qualitative data inputted into tables.
- ● **Collections:** where sets may be viewed..
  - ○ **Sets**: separate items that may be grouped together by the researcher.
- ● **Queries**: stores questions about the data.
- ● **Reports**: predefined or customized.
- ● **Folders**: groups sources together according to the preferences of the researcher.

Examples of NVivo datasets

An NVivo sample project is included in versions 9-12. See: [How to access the NVivo sample project?](#)

NVivo data are also available with limited access at the Qualitative Data Repository. See example: [Burke_EDI_NENA.NVivoProject.2017.nvp](#)

NCapture (.nvcx)

NCapture is a web browser extension that assists in gathering selected information from the internet to be transferred to an NVivo project for analysis. Submissions by researchers may contain .nvcx files that are related to the project file (.nvp, .nvpx), but not necessarily imported into it.

Key Questions to ask of the data

Which format of NVivo 12?

NVivo 12 has three formats: Mac, PC Pro and PC Plus. These formats may impact your curation workflow and considerations. NVivo for Mac will be unable to directly open a PC format. Conversions from PC to Mac formats can only be made within a PC format. There is limited functionality for certain data types and interoperability in the Mac format, so a data curator may prefer to work in a PC format. The Mac and PC formats also have different file extensions, see below.

See the QSR comparison table for NVivo 12 formats:
https://www.qsrinternational.com/nvivo/nvivo-products/nvivo-product-suite-overview

NVivo 12 for Mac (.nvpx)

This format for Apple computers will open previous versions of NVivo for Mac, versions 10 & 11. The older file will be copied to the NVivo 12 format. It will not directly open current or previous versions of NVivo for PC. These must first be converted to a Mac format from within an NVivo 12 PC format.

NVivo 12 for Mac will hide data that is not supported by this format. Hidden data may look like empty folders. Erasing these folders will delete data that may otherwise be visible in supporting formats.

NVivo for Mac doesn't support importing YouTube content. The curator will not be able to access YouTube content in submitted NCapture files (.nvcx) while working from the NVivo for Mac format.

This link will take you to a list of known issues with NVivo 12 for Mac format:
https://www.qsrinternational.com/nvivo/support-overview/faqs/are-there-any-known-issues-in-nvivo-12-mac

NVivo 12 for PC Pro (.nvp)

This format for computers using the Microsoft Windows operating system will open previous versions of NVivo for Mac or PC: NVivo 11, NVivo 10, NVivo 9, NVivo 8, NVivo 7, NVivo 2, NVivo 1, N6 (NUD*IST 6), N5 and N4. The older file will be copied to the NVivo 12 format.

NVivo 12 for PC Plus (.nvp)

This format for computers using the Microsoft Windows operating system will open previous versions of NVivo for Mac or PC: NVivo 11, NVivo 10, NVivo 9, NVivo 8, NVivo 7, NVivo 2, NVivo 1, N6 (NUD*IST 6), N5 and N4. The older file will be copied to the NVivo 12 format.

The plus version supports PAJEK software, which is used for analysis and visualization of large networks. This version also has the capability to support additional visualization types. Please see the QSR International comparison table of NVivo 12 formats for details.

This link will take you to a list of known issues with NVivo for PC formats:
https://www.qsrinternational.com/nvivo/support-overview/faqs/are-there-any-known-issues-in-nvivo-12-windows

Problems opening the file

Older Versions

Nvivo 12 formats may not be opened by previous versions of NVivo. However, older versions may be opened in NVivo 12. See details in the above sections for specific NVivo 12 formats.

NVivo 12 for Mac

Projects cannot be opened from a network drive. NVivo 12 for Mac projects that exceed 10GB might not open in NVivo for Windows.

Content issues

Data loss from conversions

Data loss may occur during conversions from one format to another. This may occur in conversions between NVivo formats, other QDAS formats, or preservation formats. Known problems converting between NVivo formats include loss of content access when converting from a PC format to a Mac format, or converting from a PC Plus format to a PC Pro format. Conversion loss between NVivo and other QDAS formats can be anticipated due to differences in feature support. REFI-QDA compliant software *should* have a pop-up box with details of content that cannot be transformed to the new format, but this automatic feature shouldn't be considered foolproof. Features, and therefore content, may also be lost in conversions to preservation formats, so it's important to discuss what is necessary for reuse with the researcher when possible.

Sensitive and restricted content

Data involving human subjects should be examined for exposed identifiers. Major content areas (i.e. sources) are obvious places to check, but other areas can also harbor identifying information. The nodes or codebook should be checked carefully, as identifiers there may be indirect. Be aware that memos, notes, annotations, journal entries and linkages may also contain identifiers. Please see the de-identification information source list in the appendix section for more help.

Data within a project may also be copyrighted or have restricted terms of use. For example, published works that have been loaded into a project may be copyrighted materials, or social media data either loaded into a project or within an NCapture file may have restricted terms. Be sure to check that sources are available for sharing and reuse. Please also see: [Social Media: Twitter Primer - expected early summer of 2020]

Another consideration is responsible and ethical sharing of information gathered from other cultures or groups. Please see: [Oral Histories Primer - expected early summer of 2020]

Key Clarifications to ask the researcher

To enclave or not to enclave? Or where are we putting this data?

Due to the sensitive nature that is usually associated with qualitative data, curators should anticipate that data sharing will likely be limited. It may be possible to share this data if certain permissions have been made, such as conditions from an IRB and consent forms. Please see [Human Subjects Essentials Primer - expected spring 2020] for more information on IRB and consent forms. A data sharing enclave may provide an ideal limited access environment to share this data.

Some examples are:
ICPSR (Virtual Data Enclave)
Qualitative Data Repository

File format preferences for reuse

Keeping the proprietary format

It may be preferable to keep the NVivo format, because components may lose some functionality with exporting and some may not export. It is ideal to work with the researcher in understanding the expected lifespan and purpose of reuse for the data. This will inform whether or not to apply practices of transformation to the format.

Exporting components

Nvivo for Mac

To export files or memos:
1. Click on the item in Detail View or List View.
2. On the Data tab, in the Export group, click Items.
3. (Optional) You can change the name of the exported file or choose to save it in a different location. If you are exporting a document, you can also change the file format.
4. Click OK.

(QSR International, n.d.)
For more information: Export Files or Memos

NVivo for PC formats

To export files or memos:
1. Click on the item in Detail View or List View.
2. On the Share tab, in the Export group, click Export.
3. (For audio, video and pictures) Select the component of the file that you want to export—for example, you could export the video component as a .wmv file.

4. (Optional) Under Save As, select the Browse button to change the name, location or format of the exported file.
5. Select any other options that you want to include in the export.
6. To display the exported file after exporting, select the Open on Export check box.
7. Click OK.

(QSR International, n.d.)

For more information: [Export Files or Memos](#)

Data content necessary for reuse

Codebook or The Nodes are the Codes

Nvivo projects contain nodes, rather than a codebook. However, a codebook (format: .docx or .xslx) might have been exported from the nodes and included as an independant file with the deposit. The choice to transform NVivo project by exporting its components to preservation formats would necessitate exporting nodes to a codebook. Before exporting, conform with the researcher that nodes are complete, appropriately named and consistent in structure, and that node descriptions are complete and consistent.

Export a codebook (NVivo 12 for Mac)

To export a codebook:

1. On the Data tab, in the Export group, click Codebook.
2. (Optional) If you have node folders in your project, you can choose which folders you want to include in the codebook.
3. Click Select.
4. You can now change the name, location or format of the exported codebook, then click OK.

(QSR International, n.d.)

For more information: [Export Nodes](#)

Export a codebook (Nvivo 12 for PC formats)

To export a codebook:

1. On the Share tab, in the Export group, click Export Codebook.
2. Under Save As, select the Browse button to specify the name, location or format of the exported file, and then click Save.
3. (Optional) If you have node folders in your project, you can choose which folders you want to include in the codebook. By default, the Automatically select subfolders check box is selected, so any subfolders will be included if you select or deselect a parent folder.
4. (Optional) Select Include number of files and references check box.

(QSR International, n.d.)

For more information: [Export Nodes](#)

Keeping the project contextual

If exporting components of NVivo to preservation or other formats, consider the potential importance of sets, linkages and memos in providing context that would assist reuse of the more explicit sources of data within a project. Be aware that sets and linkages may not export directly from a project, and linkages contained within other media may be lost in some conversions.

Description

Metadata

1. QuDEx - Qualitative Data Exchange Schema
2. DDI - Data Documentation Initiative

QuDex is an XML based metadata schema that can preserve annotations and relationships of data and other related objects.

DDI is a metadata standard for social science data that can be used in conjunction with QuDEx.

See descriptions and examples at UK Data Archive: [Metadata Standards](#)

ReadMe

A ReadMe file may serve to explain how various components within an NVivo project, or those exported from an Nvivo project, relate within the context of the original research. The ReadMe file can also be used to describe how components may be accessed; such as, how to view the contents of an NCapture file, if one is included in the submission. It can also describe conversions between formats of NVivo or other transformations that may affect opening or accessing the contents of the file.

Cornell guide to writing a README: https://data.research.cornell.edu/content/readme

Preservation actions

Nvivo project files are a proprietary format. According to Van den Eynden and Chatsiou (UK Data Archive, 2011), "It is preferable for long-term preservation to archive individual data and documentation files, rather than the entire NVivo project file."

(Note: Preferred formats may differ by the repository or enclave, so be sure to inquire about specific preferences for deposit.)

The UK Data Archive provides this table in their guide *DATA MANAGEMENT FOR QUALITATIVE DATA USING NVIVO9*, authored by Van den Eynden and Chatsiou:

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	

Retrieved from: https://ukdataservice.ac.uk/media/622387/ukda-datamanagement-nvivo.pdf

Further considerations

Reuse across QDAS software

The REFI-QDA standard is an open standard allowing data to be reused across Qualitative Data Analysis Software. Data is exported from one program and imported into another if both QDAS programs have implemented this standard; see tables for project and codebook. Be aware that data loss may occur in this process. For more information: Blommestein (2019)

All three versions of NVivo 12 are compliant with the REFI-QDA standard.

Workflow based on the Data Curation Network CURATED steps

Some NVivo specific workflow considerations are listed below. Other considerations of the data may apply. See the original workflow checklist: Data Curation Network Checklist for DCN Curators

Check

- Check for NVivo project to open and run as expected.
- Check components of the project for completeness. Keep in mind the NVivo project will contain multiple media types and implicit contextual associations.
- Check components of NVivo project for human subjects information, copyrighted sources, or sources that may violate terms of use if shared.
- Were there any NCapture files submitted? Are these already imported into the NVivo project? Do these independently contain human subjects information, copyrighted sources, or sources that may violate terms of use if shared?
- Has documentation been provided by the depositor that would allow reuse of any sensitive or restricted content associated with this submission?
- Has a codebook been separately provided? Are the nodes within the NVivo project the only codes?

Understand

- Examine components of the NVivo project for human subject identifiers.
- Confirm that the documentations/permissions for sharing sensitive or protected information are appropriate for the requested level of access to the data.
- Reach out to the researcher to understand what file format(s) or components are necessary for the intended reuse of the data. Be prepared to discuss any legal or ethical restrictions, and to assist understanding options for sharing, such as: what components may be shared, the pros/cons of file format options, and limited/restricted access options for publishing the data.

Request

- Request any additional documentation or permissions for human subjects, copyrighted sources, or otherwise protected sources that are needed to support the intended level of access to the data.
- See the [Human Subjects Essentials Primer - expected in spring 2020] for more on necessary documentation.
- Request any missing information necessary for reuse.

Augment
- See: Checklist for DCN Curators for information about metadata

Transform
- A codebook may be exported, though it should be checked for completeness.
- Sources may also be exported to preservation formats. Context from less obvious places in the project (e.g. sets, links) may also be necessary to preserve in some way.
- Be aware that transformations using REFI-QDA standard, or other transformations to a preservation format may result in data loss.

Evaluate
- Expect the NVivo project, its exported components or NCapture files to be limited/restricted access due to containing sensitive information. FAIR principles may not necessarily be the primary concern of the format choice for restricted data. Intended users or policies within an enclave or restriction may prefer certain format choices, such as keeping a proprietary format. However, the project description and a README file can be made open (devoid of human subject identifiers) and adhere to FAIR principles.
- See: Checklist for DCN Curators for information about FAIR

Document
- Remember to record actions taken during the curation process.

Bibliography

Bazeley, P., & Jackson, K. (2014). *Qualitative Data Analysis with NVIVO (ed.)*

Appendix
NVivo versions

Extended List of NVivo Versions			
NVivo for Mac		NVivo for PC	
Year	Version	Year	Version
2018	12 See: NVivo 12 for Mac	2018	12 See: NVivo 12 for PC
2015	11	2015	11 - Pro and Plus editions added
2014	10 - First version NCapture - available for Mac	2012	10 NCapture - First version
		2010	9

		2008	8
		2006	7
		1981-2005	See: [Our History - QSR International](#)

Additional Information:
- [NCapture Help](#)
- [NVivo 12 Feature Comparison](#) (Mac; PC Pro and Plus)

Export data information list
The following links will take you to QSR International's help pages for NVivo 12:
- NVivo 12 for PC: [Export sources](#)
- NVivo 12 for Mac: [Export sources](#)

The following link will take you to Uk Data Archive information for exporting data and documentation in NVivo 9 for the purpose of archiving:
- [Chatsiou (2011)](#)

List of supplementary NVivo primers
The following are additional NVivo primers that explore aspects of the format in more depth:
- NVivo (Name to be determined) - expected early summer of 2020

De-identification information source list
The following links will take you to information sources for the de-identification of data involving human subjects:

Human Subjects Essentials Primer - expected early spring of 2020

Protecting Identifiers in Human Subjects Data: https://guides.library.jhu.edu/protecting_identifiers

Methods for De-identification of PHI:
https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html