Digital literacy training

NVivo 12 Pro Intermediate

More Tools for Coding & Project Management
2022
# Table of Contents

**NVivo Intermediate: Project Set-Up & Coding** ................................................................. 1
  Opening NVivo and a Sample Project ................................................................. 1
  Mac vs PC Versions of NVivo ............................................................................. 1
  Compatibility with other versions of NVivo ................................................... 1

**Other Methods for Creating Nodes & Cases** ...................................................... 2
  Mind Maps ............................................................................................................. 2
  Auto Code ............................................................................................................. 2
  Auto Coding Datasets ......................................................................................... 3
  Auto Coding Word Documents ........................................................................... 3
  Auto Coding Cases ............................................................................................. 4
  Auto Coding Nodes ............................................................................................ 4
  Auto Coding When Importing Survey Spreadsheets ........................................ 4

**Creating & Coding to Relationships** ................................................................... 5

**Deciding on a Coding Scheme: Word Frequency Queries** ..................................... 6
  Word Frequency Results ...................................................................................... 7
  Excluding Irrelevant Words from Searches ....................................................... 7
  Word Clouds ......................................................................................................... 8
  Tree Map .............................................................................................................. 8
  Cluster Analysis .................................................................................................. 8
  Looking at the Results in Context ...................................................................... 9

**Initial Mass Coding: Using Text Search Queries** ................................................ 9
  Text Query Results Preview ............................................................................... 10
  Saving the Query Preview Results as a Node .................................................... 10
  Expanding the Coding Context of References in Results or a Node ................. 10
  Automatically Expanding the Coding of References in a Node ...................... 11
  Word Trees ......................................................................................................... 11
  Querying Within Project Items ......................................................................... 12

**Checking Inter-Coder Reliability** ...................................................................... 12
  Exporting the Codebook ..................................................................................... 12
  Creating a Set Folder ......................................................................................... 13
  Creating a Project Copy ..................................................................................... 13
  Coding Comparison Query ................................................................................ 14

**Exporting** .............................................................................................................. 15
  Exporting Results of Queries ............................................................................. 15
  Exporting Nodes ................................................................................................ 15
  Exporting the List View of the Nodes Area ....................................................... 15
  Exporting Classification Sheets ......................................................................... 15

**Other resources** .................................................................................................. 16
NVivo Intermediate: Project Set-Up & Coding

NVivo can do a large number of things to help you analyse your qualitative data. This course will show you how to use the essential tools available in NVivo for qualitative data analysis in an academic context. For further information about NVivo’s full capabilities, please refer to the NVivo 12 help website:

PC: help-nv.qsrinternational.com/12/win/v12.1.110-d3ea61/Content/welcome.htm
Mac: help-nv.qsrinternational.com/12/mac/v12.1.110-d3ea61/Content/welcome.htm

This session will cover how to:

- use Word Frequency Queries to decide on a coding scheme
- initially mass code your text data using Text Search Queries
- Auto Code to create contextual Nodes and Cases
- check the reliability of your coding scheme using Coding Comparison Queries

Opening NVivo and a Sample Project

When you open NVivo it will ask you to create a User ID: enter your name and initials (must be unique to the Project). This is very important when you’re collaborating with others. At the top of the Start Screen there is an option to open a new Blank Project or a new Sample Project and on the left-hand side there will be a list of Recent Projects opened. For this session open the Sample Project which is designed to help you explore how NVivo can be used to analyse your own data. It is a two-year study (2008-2009) documenting community perceptions of development and land-use change on coastal communities in the Down East area of Carteret County, North Carolina, USA.

Mac vs PC Versions of NVivo

- If you are collaborating with others on the same Project, it’s best to use the same version of NVivo (e.g., NVivo 12).
- Mac versions of NVivo do not have the same capabilities as Windows versions (e.g., you cannot Create Cases at import on Macs). Windows Projects can be converted to Mac and vice-versa. If you need to use these functions, save a copy of your Mac Project in a Windows format, use a PC to do what you need to, then convert it back to Mac format so you can continue working on your Apple computer.
- These notes apply to the PC version of NVivo, so if you are working on a Mac and have trouble finding what you need, please refer to the online NVivo 11 Help for Mac:
  help-nv.qsrinternational.com/12/mac/v12.1.110-d3ea61/Content/welcome.htm

Compatibility with other versions of NVivo

Projects created in older versions of NVivo can be opened in newer versions and vice versa (e.g., NVivo 12 Project in NVivo 12), with the exception of “NVivo Release 1”. If you open an NVivo 11 or 12 Project in "Release 1", the Project will be converted to that newer format, and once it is saved you cannot open it in an older version again.
Other Methods for Creating Nodes & Cases

Mind Maps

NVivo Mind Maps help you to create a diagram of your coding scheme and its hierarchical structure at any point in the Project. However, if you begin with a Mind Map first, you can convert it into actual Nodes in the Codes area with the click of a button.

On the Explore tab click Mind Map. Enter a name for the map in the Name box, and an (optional) Description, then click OK. An initial shape is already prepared for you to enter a label for the main theme of the map (e.g., Research Project Title). The child ideas from this main shape will essentially be the parent Nodes of your coding scheme (otherwise you need to have one map for each parent Node which is inefficient).

If you want to add a “child” idea (e.g., a shape connected to the main idea) click on the shape and on the Mind Map Tools tab, click on Child idea (alternatively right-click on the shape and select Child idea from the menu). Then give it a label (e.g., Theme 1). If you need to relabel it, just double-click on the word to edit. You can add a Sibling idea (a shape at the same level in the hierarchy) in the same way.

If you want to create an unconnected shape that will become a free Node in your Nodes folder, click Floating idea on the Mind Maps Tools tab. Note: you will not be able to add siblings or children to the shape.

You can change the colour of a shape by clicking on it and changing the colour from the Fill drop-down menu on the Mind Maps Tools tab. You can also change the layout, size, and text formatting on this tab. When you’re happy with your Coding structure, select Create as Nodes or Cases from the Mind Maps Tools tab, click on the Nodes folder and then OK. They will then appear in the Node folder of the Codes area.

Note: When you have created your Nodes from the map, I recommend rearranging the new Nodes to make the parent Nodes (which will appear as children under your main theme – shown in purple below) true parents by cutting and pasting them in the blank space in the list (or copying and pasting them into the main Codes folder). Then you can remove the superfluous project theme Node.

Auto Code

You can also create Contextual Nodes (such as a Node for each question in an interview/survey so you can analyse the responses separately) and Case Nodes using the Auto Code function.
Auto Coding Datasets

This works best if you have a spreadsheet Data file where each row represents responses from one case, and each column represents a Node (such as a survey question). Then you can automatically code datasets by either:

- **rows** to create a Case for each respondent and code the responses in that row to that Case, or
- **columns** to create a Node for each question in a survey and code the responses in the column to that Node.

Example: Auto Code the columns in the Survey Responses Data file as question Nodes.

In the List View of the Data area, select on the dataset(s) you want to Auto Code (e.g., the Survey Responses Data in the Survey folder). On the Home tab click **Auto Code** and the Wizard will begin to take you through the process. Click **Code at nodes or cases for selected columns** then Next. Move the appropriate column(s) you wish to Auto Code into the Selected Columns box by selecting them and clicking the right arrow button (e.g., all of the available columns) then click Next. Select a location for the Node (e.g., under New Node in the location Nodes), and give the new Node a name (e.g., Survey Questions). Then click Finish. This will create a new parent Node called **Survey Questions** with multiple child Nodes in the Nodes folder.

Example: Auto Code the rows in the Survey Responses Data file as Case Nodes.

Similar to the procedure above, only when the Wizard opens, select **Code at nodes or cases for selected rows**. Choose the column that contains nodes or case names from the drop-down folder (e.g., Respondent). You also have the option of filtering out some cases (e.g., cases below a certain age in the age column) or grouping them (grouping Cases aged 18-30, 31-49, etc. if they aren’t already), by clicking the Filter and Group button next to this menu. This may be useful if your dataset isn’t already “cleaned”, however it is **best practice to clean the data before importing it** in a more suitable program, such as Excel or SPSS.

Your data is unlikely to warrant a hierarchical structure, so click Next. Select the columns in the data file which you want to code to the Case Nodes in the Available columns box by holding the CTRL key and clicking on them (e.g., all question columns), then move them into the Selected columns box by clicking on the right arrow, then click Next. Usually, you won’t already have Cases for this new import, so from the Under drop-down menu, select New Node or Case (if you are adding content to existing Cases, select Existing Node of Case). The Location box should have Cases already selected, however if you are using Auto Code by Rows to create Nodes, change this to the Nodes folder. The Name of the Cases is Survey Responses by default, but you can change this to something more appropriate, then click Finish. The new Cases will appear in the higher-order Cases Folder, in the Cases Area.

Auto Coding Word Documents

**Auto Code** can also be used in Word documents, if you have used Heading Styles to define content relating to Cases, questions asked, general themes etc. In the Interview Transcripts of the Sample Project, they have used Heading 1 for each interview question asked and Heading 2 for each speaker’s name (see excerpt below). Everything the speakers say in the transcript is in Normal text.

Q1. Connection to Down East

**Henry**

Tell me about your personal and family history in Down East. How long have you or your family been living Down East full time or part time?

**Barbara**

My family moved here when I was two years old in 1939. My parents still live here. They live down in Gloucester. My father was raised in Massachusetts, in town, and went to Brandeis Elementary and middle school and high school, then moved away for college. We lived here most of my life although we moved away.

**Henry**

And you’ve lived Down East how long?

**Barbara**

Auto Coding Cases

It’s probably easier to explain how Auto Coding works using the interview transcript above as an example. If you want to create a Case for each speaker, you will need to code by **Heading Style 2**. NVivo will automatically create a new Node for each unique title in that style (e.g., one Case Node called "Henry", another called "Barbara"). It will then automatically identify all of the sections of text belonging to Henry and code it to the **Henry Case**, and all of the text belonging to Barbara to the **Barbara Case**. It does this by identifying all of text in between instances of **Heading Style 2** (regardless of any other **Heading Style** used in between) and coding it to the Case Name written in the preceding **Heading Style 2**. If NVivo finds the identical text in **Heading Style 2** again (e.g., indicating more text relating to "Barbara"), it just adds that section of text to the existing Case Node it created earlier.

Example: Auto Code to Case Nodes in the **Interview** Word Documents.

Ensure all the documents you wish to Auto Code are identically formatted using consistent heading names and styles. If you find any inconsistencies, it’s easier to correct them now before coding. Select all the documents you want to code (e.g., all the text documents in the **Interviews** folder) in the Data area using the CTRL key to click on multiple files. On the Home tab, click on **Auto Code** and select **Paragraph Style**, then select the appropriate paragraph style(s) from the list (e.g., Heading 2 to create a Case Node for each speaker). Click **Next** and then in the **Under** box select **New Node** (unless you are adding information to existing ones) and then click on the **Select** button to change the default Location from **Nodes to Cases**. Give the **group** of Cases a **Name** (e.g., Interviewees), then click **Finish**. You’ll find the new Case Nodes in the Cases folder in the Cases Area.

Auto Coding Nodes

If you want to create a contextual **Node** for responses to each question asked in this interview transcript, you will code by **Heading Style 1**. Again, it will follow the same procedure as indicated above. **Please Note** that if there is a typo in your **Headings**, NVivo will create a duplicate Node for that Node/Case, and this can simply be merged into the primary Node (without the typo) as described in the **Introductory Course Notes**.

Example: Auto Code to Question Nodes in the **Interview** Word Documents.

Follow the same procedure as for Auto Coding Cases but leave the default **Location: Under New Node or Case** set to **Nodes**. Give the **group** of Nodes a **Name** (e.g., Questions), then click **Finish**. You’ll notice in the Nodes folder there are 2 Nodes for Q.6. This is because the questions were not identical, so you will need to merge these two together as described in the NVivo Introduction handout.

**Note:** You can also create a hierarchical coding structure for Nodes if you select to Auto Code by multiple Heading Styles at the same time (e.g., **Heading Style 1** for the parent Node, and **Heading Style 2** for the child Nodes). Be sure to enter them into the paragraph styles box in the correct parent/child order. You can also Auto Code documents by **Speaker** if you have identified the person speaking in the transcript at the start of each paragraph of the transcript (not using **Heading Styles**). This is a more laborious process however, as you need to type in the names of every speaker in the documents first so NVivo can identify them. So, check for typos in people’s names beforehand.

Auto Coding When Importing Survey Spreadsheets

If you have a survey spreadsheet from an online survey application, such as Survey Monkey, you will already have all the information you need about your Cases and Questions neatly organised. Luckily, you can create **Case Nodes, Question Nodes, Case Classification Attributes** and populate them with their relevant details upon import, in one go. The process is easiest if you follow these rules for your spreadsheet set-up:

- **The first row of the spreadsheet** must contain the **Heading** for each column. You can have Headings in the first 2 rows, but there are no other options.
The first column should be reserved for your **Case Identifier** (e.g., name, ID, etc.) and there should be **NO duplicates**.

The **Case Attributes** (e.g., demographics like age, etc.) will have their own columns, and make sure the values are consistent and there are no typos before import. NVivo is case-sensitive, so if you have "Male" written with a capital M in some cells and a lowercase m in others, these will be seen as two different Attribute values).

Each row must represent one **Case**. If you do have duplicates, you could merge them in NVivo later, but it’s best to deal with the problem before import.

Do not have superfluous information like annotations in the spreadsheet before import, or these will pollute your data and be harder to remove in NVivo.

On the **Import** tab click on the file format your survey is in (e.g., Excel) or right-click in the Detail View of the **Files** Area and select **Import Survey** and choose the format from the menu. Select your survey file (Note: you may need to change the file format type from the drop-down menu next to **File Name** if you don’t see it), then click **Open**. The Survey Import Wizard will guide you through the process, providing previews of the result of your choice throughout. Click **Next** on the first screen, then at Step 2 tell NVivo how many of the **first rows** of your spreadsheet contain headings only (best practice is only 1, but you can have 2 if you select it from the drop-down menu). If you have any dates in the file, select the date format used from the drop-down menu, then click **Next**. Choose where you want the Cases to be located (the default is usually fine), select the column containing the Case identifier from the drop-down menu, and choose the Case Classification you want to use for the Cases in the spreadsheet (usually **Create new classification**, but you can add to an existing one if need be). Click **Next** and check that NVivo has correctly identified the purpose of each column: Attributes should be **Closed Ended**, Questions should be **Open Ended** and if there are any superfluous columns in the spreadsheet you can choose not to import them at this stage too. Then click **Finish**.

**Creating & Coding to Relationships**

You can create Relationships between Cases and/or other project items, and you will see these relationships in the **Relationships** folder in the **Codes** area. Relationships are technically a type of Node which can be used in two ways:

- to note general relationships between Project Items (e.g., the **Person Case Margaret** “is a member of” the **Place Case Cartaret Catch**),

- to collect evidence for a relationship between **Nodes** (e.g., coding data to support a theory that **Community Change** “contributes to” the **Environmental Change**).

Relationship types (e.g., “is a member of” and “contributes to”) need to be created FIRST in the **Codes** Area. Then these can be applied to create Relationships between Project items later.

To create a **Relationship Type**, right click on the **Relationship Types** folder in the **Codes** Area and select **New Relationship Type**. Type in a name (e.g. “negative impact on”) and [optional] description. Then select the appropriate direction (e.g. *one-way*).

To create a **Relationship between Items** (e.g., **Infrastructure** “has a negative impact on” **Habitat**) right-click on the **Relationships** folder in the **Codes** Area then select **New Relationship** (or select **Relationship** on the **Create** tab). Next to the **From** box, click on **Select** button and choose a Project item (e.g., **Infrastructure Node**) and click **OK**. Then choose a **Relationship Type** from the drop-down menu (e.g., **negative impact on**) and selecting another Project item in the **To** box (e.g., **Habitat**). Then click **OK**, and **OK** again.

To code **content to a Relationship** (coding evidence to support that Relationship), select the text you want to code, right-click on the selection and choose **Code**. Expand the **Relationships** list by clicking on the triangle next to it and select the Relationship you want to code to (**Infrastructure** (negative impact on) **Habitat**), then click **OK**. You can open the
Relationship “Node” to see all the relevant information coded to it, by going to the Relationships folder and double-clicking on the relevant Relationship Node.

**Note:** Relationships between Cases can be used in NVivo 12 Plus and NVivo Release 1 to perform an approximation of Social Network Analysis (SNA). However, if you want to perform a SNA with a large amount of data in a statistically appropriate way, it is best to use a program more suited for the purpose. The CSIRO has a free open-source tool called **Constellation**: www.constellation-app.com/index.html and another one called **StellarGraph**: www.stellargraph.io/ which you may find helpful.

### Deciding on a Coding Scheme: Word Frequency Queries

In the initial stages when you are still deciding on a coding scheme, a **Word Frequency Query** could be useful. This will produce a list of all the words that occur most often in the text documents you have (including memos, transcripts etc. unless you specify otherwise). You can search the entire content of your qualitative data, or you can narrow it down to specific groups of people or individual questions, etc.

**Example:** **Word Frequency Query for Interview Question 4.**

On the **Explore** tab, click on **Word Frequency**. This will open a new **Detail View** with the Word Frequency options. Choose where to search in the **Search in** area. For this example, click **Selected Items** and then click the + next to the **Nodes** folder to see the other folders within it. Then click on the name of the **Autocoded Responses** folder (not the checkbox - if you do all the Nodes within will be selected) and then on the + next to the **Autocoded Interview Questions** folder and tick the box for **Question 4**. Then decide if you want to **limit the number of results** (usually the default of the 1000 most frequent is fine) or change the largest **letter-length** of the words you want in the results (the default of 3 letters is fine for English, so you don’t get words like “to”, “a”, “an” or “I”). In the **Grouping Area**, decide on the grouping criteria by sliding the indicator (e.g., I recommend looking at results from both **With stemmed words** and **With synonyms** Queries - see table below for descriptions of all 5).

**Note:** The higher the similarity, the longer it will take to run.

**Note:** A **Text/Word Search Query** does not search text within PDFs created by scanning paper documents (each page is a single image). Consider using optical character recognition (OCR) to convert the scanned images to text before importing.

**Note:** Only words in the transcript fields are included in Queries when searching audio and video Data Files and only in "codable" fields in datasets.

### The five Text Match levels:

<table>
<thead>
<tr>
<th>Grouping Level</th>
<th>Returns</th>
<th>Example (sport)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exact matches</td>
<td>Exact matches only</td>
<td>sport</td>
</tr>
<tr>
<td>With stemmed words</td>
<td>+ Words with same stem</td>
<td>sport, sporting</td>
</tr>
<tr>
<td>With synonyms</td>
<td>+ Synonyms</td>
<td>sport, sporting, play, fun</td>
</tr>
<tr>
<td>With specializations</td>
<td>+ Specializations</td>
<td>sport, sporting, play, fun, running, basketball</td>
</tr>
<tr>
<td>With generalizations</td>
<td>+ Generalizations</td>
<td>sport, sporting, play, fun, running, basketball, recreation, business</td>
</tr>
</tbody>
</table>
It’s good practice to save a detailed record of any analyses you conduct in NVivo or otherwise. This serves as both a reminder as to what you did and how you did it, and also provides a short-cut to run the analysis again if you need to. To save the Word Frequency Query details, click **Add to Project** and give the Query a name, then select **OK**. Then click **Run Query**. If you need to run the same Query again later, you will find it in the **Queries** folder in the **Search Area**. Just right-click on the Query and select **Run Query**.

**HANDY HINT:** If you need to **run multiple similar Queries** (e.g., Word Frequency for each interview question with the same grouping level etc.), **copy** your first Query (e.g., **Frequency Q1**), then **paste** it into the blank space in the **Queries** List View. This will create a new saved **Query** with identical parameters, called **Frequency Q1 (2)** and you will need to rename it to reflect the changes you will make (e.g., **Frequency Q2**). Right-click on this copy and select **Open Query**. Then change the parameter(s) you wish (e.g., **Selected Items** to Q2), and click **Run Query** to make the changes “stick” in this newly created and saved Query.

**Please Note** that if you ever change a parameter in a saved Query, then run it, that parameter is permanently changed.

### Word Frequency Results

There are several tabs on the right-hand side of the **Detail View** which present the results in different ways. From these you can decide on an appropriate coding scheme and even begin to create Nodes. The **Summary** tab shows you a list of all the most frequent words, their frequency **Count**, other summary data and if you have chosen a **Grouping** level higher than **Exact matches**, you will also see the other words it has included as belonging to each Word Count (e.g., the word count for Change also includes that for Changed, Changes, Changing).

There are also tabs for graphical representations of your Query: **Word Cloud**, **Tree Map** and **Cluster Analysis** (the latter only appears if there is enough data to run it).

### Excluding Irrelevant Words from Searches

In the results you might see words that you do not want to include in the Query search in future, such as “any”. Right-Click on the word in the results and select **Add to Stop Words List**. If you have included more than exact matches in the text match level section (i.e., >level 1), you must remove any words in the box which you still wish to be included in future searches. When you are ready, click **OK**. You can **edit** the **Stop Words List** in **Project Properties** at any time: on the **File** tab under **Info** select **Project Properties** and on the **General** tab click the **Stop Word List** button and edit the words to exclude.

**Note:** Every time you run a Query from now on, these words will not be included in Word Frequency or Text Search Query results – even if you specifically search for them. So, it is a good idea to check the **Stop Words List** before you run any Text Searches or Word Frequency Queries. There may be some words on the list that you don’t want NVivo to ignore in these Queries. You will have to remove these words from the Stop Words List, before conducting your Queries.
Word Clouds

The **Word Cloud** tab displays up to 100 of the most frequent words in your text data, varying in font size: the largest are the *most* frequently occurring, and the smallest the *least* frequent.

Tree Map

The **Hierarchy Map** tab displays up to 100 of the most frequent words in rectangles of varying sizes: the largest rectangles are for the most frequently occurring words, and the smallest for the least.

Cluster Analysis

The **Cluster Analysis** tab displays up to 100 of the most frequent words in a tree-diagram describing the co-occurrence between these words. The closer the words are on the branches of the tree map, the more frequently they co-occur. This can help by providing more context. Please note that a tree map can only be produced if there is sufficient data.
Looking at the Results in Context

If you think that one of the words in the results might be a good Node to have in your Coding Scheme, you can **view each word in the results in context** by automatically running a Text Search Query for that word or group of words. Double-click on the word in the Summary table/Word Cloud/Tree Map, or right-click and Choose **Run Text Search Query**. This runs a Text Search for the word (for exact matches) or words (for grouped matches) and opens a new tab in the **Detail View** containing all the relevant references containing the word(s).

**Note:** You can also **create initial Nodes** from Word Frequency Queries by right-clicking on the word in the results and selecting **Create as Code**. However, it is probably best to run the **Text Search Query** first, before deciding the results would make a good initial Node at all (procedure described in the section below). Of course, the initial Nodes will have to be "cleaned up" afterwards to both remove irrelevant and add relevant information.

### Initial Mass Coding:
**Using Text Search Queries**

**Text Search Queries** can be used to help start the coding process for a large amount of textual data by finding instances of text matching specific **search terms** based on your pre-defined coding scheme (**internal text data files ONLY**). Information coded to Nodes this way can always be modified later through uncoding if necessary.

**Note:** Consider checking the spelling before running Queries. You need to open a document and click the blue **Click to edit** text at the top of it. Then you will find the **Spelling** button on the far right of the **Edit** tab’s menu in the **Document Tools** group of tabs which appear.

**Example:** Create an initial, automatically coded Node for Water Quality.

On the **Explore** tab, click on **Text Search** and this will open up a new **Detail View** window with the Query options at the top. Choose where to **Search in** (e.g., Selected Folders) and tick the boxes for the folders you want to search in (e.g., in **Interviews** and **Surveys**), then click **OK**. In the **Search for** area, enter the word(s)/phrase you want to look for. You can create more complex Queries using the special features displayed when you click the **Special** button:

- **Wildcards:** If you want to search for words that have alternative spellings (e.g., colour/color), or are commonly confused with similar words (compliment/complement), Wildcards can help to find every possibility. The Wildcard * is used to stand in for one or more missing characters (e.g., g*t will find get, great, gradient etc.). The Wildcard ? is used to stand in for a single missing character (e.g., g?t will find get, gut, got etc.).

- **Fuzzy Search:** Also useful to pick up misspellings and find words with similar stems. The symbol ~ behind a word will look for similar words (e.g., analyze~ will find analyzes, analyse, etc.).

**Note:** If you use the above functions you will be limited to **Exact match only**.

- **Boolean Operators:** When you want to find a specific concept you can use a combination of words and/or excluding a word(s) in your search.
  - **AND** (or &) will search for BOTH terms (water AND quality).
  - **OR** (or |) will search for EITHER of the terms (water OR lake OR sea OR river).
  - **NOT** (- or !) will search for the first term and EXCLUDE any results that also contain the second (water NOT tap, water -tap).
  - **REQUIRED** (+), will search for items containing the required term and the second term is optionally included in results if and only if it occurs together with the first term (+Water river).
  - **Double Quotes** “ ” should be used to enclose a group of words to search for a specific phrase. If you search for a phrase, the results will not include words with
the same stem. For example, if you search for "alternative energy", you will not be able find "alternative renewable energies".

- **Round Brackets ( )** can be used to group clauses to form sub Queries.

- **Near**: If you want to find words that co-occur but you want to ensure they are mentioned in the same sentence rather than several sentences away, you can specify the distance either side of the required word (e.g., "water AND quality"~20) will find instances of water and quality within 20 words of each other only.

**Note**: the syntax for a Near Search reuses text characters in counter-intuitive ways. You do need to enclose the search terms in double quotes before adding the ~ for a Near Search as shown in the example. NVivo recognises that in this context, the quotations marks do not signify an Exact Phrase Search and the tilde (~) does not signify a Fuzzy Search.

For this example, enter "water AND quality"~20. **Spread Coding to a Broad Context**, so you can see enough information when you review the success of initial coding. Decide on the matching criteria (e.g., With stemmed words). Click Add to Project if you want to save the Query parameters and make it easier to rerun it later if you obtain new data. Then click on the drop-down menu for Run Query.

**Text Query Results Preview**

In the preview of the Query results, there will be tabs on the right of the Detail View for:

- **Summary**: A list of all the Data Files that contain the word/phrase results.
- **Reference**: The results are opened as a Node preview and the word or phrase is shown in context (which can be broadened or narrowed as outlined below).
- **Text, Picture, Audio, Video, or Dataset**: Displays the results found in each type of Data File (only relevant tabs are available).

**Saving the Query Preview Results as a Node**

In general, if you run any Query and decide retrospectively that the results in the preview would be valuable to save as initial coding for a Node, click on the Query results in the Detail View, then on the Node Tools tab click the Store Query Results drop-down menu and select Store Query Results again. From the Option list, choose either to:

- create the results as a new Node or Case, or
- merge the results into an Existing Node or Case (useful when “chunking” searches to minimise typos etc.).

Another method is to right-click somewhere inside the Query results in the Detail View and select Store Query Results. Change the location of the Node from the default Query Results to the Nodes folder (which is where you would usually want it stored), by clicking on the Select button. Give the new Node a Name and optional Description, then click OK.

**Note**: NVivo will just code the word(s) found in that Query unless you expand the coding (see how to do this below).

**Expanding the Coding Context of References in Results or a Node**

When you run a Query, by default the context for the references in the results is set to Narrow, and thus displays 5 words on either side of the result greyed-out. When you save the results as a Node, only the black text is coded to the Node, and you will need to add to this. In order to see more context to help with coding more data, you can expand the Coding Context for one reference at a time, or all references together.
To expand the context of one reference, within the Query Results or Node select and right-click on the chosen reference in the Detail View, then select Coding Context and choose the level you want:

- **Narrow** [default] to see 5 words (text), 5 percent (pictures), or hear/see 5 seconds (audio/video) more on either side of the coding reference, or (spreadsheets) see all the text in the cell.
- **Broad** to see the surrounding paragraph (text), cell or row (spreadsheets), or hear/see 20 seconds (audio/video) or 20 percent (pictures) more on either side of the coding reference.
- **Custom** to define specific reach settings.
- **Entire File**.

Then click OK. More context will appear greyed-out on either side of the coded text. If you want to expand the coding context for the entire Node, select all the text at once either using the mouse or holding down the Ctrl and A keys, then right-click on the selection and follow the instructions above.

**Automatically Expanding the Coding of References in a Node**

Sometimes it is easier to expand the coding itself as well as the context, and then remove the extra content that does not belong in the Node. To do this, when you are in the Node you can select and right-click on the chosen reference(s) in the Detail View and select Spread Coding then choose the level you want (same levels as listed for Coding Context). Again, you can do this for the entire Node as explained above.

**Word Trees**

**Word Tree** tabs are produced when you run Text Search Queries (Note: only if you choose NOT to spread coding or save the results). The branches of the **Word Tree** represent the various contexts in which the word or phrase occurs. It can be useful for finding recurring themes in phrases that surround the word.

You can click one branch to highlight the other end of the branch to see the actual context. When on the **Word Tree** tab, a main **Word Tree** tab appears at the top of the window. This allows you to change the **Root Term** (e.g. to the quality of the water), the **Branch Order** (e.g., by **Number of Matches**) and **Context** level (the number of words surrounding the root term).
You can search for an exact quote from the **Word Tree** in all of your data by right-clicking on the appropriate branch, then selecting **Run Text Search Query**. This has the benefit of finding the quote quickly and getting a link to open the **Data File** from which it comes so that you can see it in more context.

**Querying Within Project Items**

NVivo also has a handy shortcut which allows you to Query within specific Project Items that are open (e.g., Nodes, Cases, Documents, etc.). For instance, when you have a **Node** open, on the **Node Tools** tab there is an icon to **Query This Node** with the following options:

- **Word Frequency Query in This Node** (automatically runs a standard Query whose parameters can be changed and run again).
- **Text Search Query in This Node** (automatically selects that specific Node to search in, and you can change the parameters before its run).
- **Nodes Coded at This Node** (runs a **Group Query** to identify all the other Nodes coding the same Node content).
- **Cases Coded at This Node** (runs a **Group Query** to identify all the Cases associated with the Node content).

**Checking Inter-Coder Reliability**

An important measure of the reliability (consistency) of your coding scheme is to see how well two Coders agree with each other regarding the application of the coding scheme. **Checking inter-Coder reliability is important to do at the beginning of the coding process**, so that if any changes need to be made to your coding scheme you don't have to recode the whole Project. It involves asking a second person to code a small sample of your data also, and then seeing how much overlap there is with your own coding. If overlap is high, then reliability of your coding scheme is high. Then the readers of your research can be more confident that if they were also to apply your Coding Scheme to your data, they would come up with similar results.

Checking reliability involves finding at least one other person familiar with your research area and providing them with your NVivo **Codebook** or your detailed coding scheme in another format.

**Exporting the Codebook**

As mentioned in the Introduction to NVivo course, it is good practice to add a detailed description for each Thematic Node you create in your Coding Scheme. This serves as your Codebook outlining the definitions for each Theme, rules for applying it to the data (coding) and, if possible, examples of text that would be coded in such a way. If you have not already gone through this process in NVivo (or elsewhere), you can add a description for each Node by right-clicking on a **Node** in the **Nodes Area** and selecting **Node Properties**. Enter your coding rules in the **Description** box and click **OK**.

To export your NVivo Codebook, go to the **Share** tab and click on **Export Codebook**. Select the relevant folder containing your Thematic Nodes by ticking the appropriate box (usually just the Nodes folder). You can also choose to include the number of files and references coded to each Node, by ticking the appropriate box at the bottom, but this would only be needed at the end of the Project when coding is finished. Then click **OK**. A **Word** file is produced and saved showing a formatted table like the one below for the **Sample Project**.
Creating a Set Folder

Select a small random sample of your Data for your second Coder to code. It is helpful for them if you create shortcuts to only these files in a Sets "meta-folder". To create a new Set folder, go to the Search Area and right-click on the higher-order Sets folder. Select New Set... and type in a name (i.e., Interviews for coding comparison) and description (optional). Then click OK. Now you have an empty folder to populate with links to any type of Project items (e.g., Files, Cases, Externals, Memos). To add new links, click on the relevant Set folder and right-click in the blank-space of the List View. Select Add Set Members and click on the + next to the Files folder, click on the name of the folder you want to add items from (e.g., Interviews), then tick the boxes for the files you wish to add to the Set. Then click OK. To remove items from a set, select the items you want to remove and on the Home tab click Delete (or right-click on the item and select Delete). Now all the second Coder needs to do is look in the folder to see which files they need to code.

To make the process even easier for the second Coder (particularly if you have a large and complex coding scheme), you can also limit the number of Nodes they need to code the data to and put links to these in another Set folder as a reminder. Prioritise those that are particularly difficult to define or potentially ambiguous, as they are more prone to unreliable coding. Remind your second Coder that they only need to code to the Child Nodes and not their parents as well, because you can always turn aggregation on if need be. If you cannot find a 2nd person, code this sample of Data yourself a second time after a few weeks using a different User Profile. Then conduct an inter-Coder reliability test to see how consistently you coded both times.

Creating a Project Copy

It is best to give your second Coder a copy of your Project to code and NOT your Master Copy. It is also best to perform the Coding reliability check in this copy of the Project, rather than importing that person’s coding into your Master Project (details on how this is done in the next section under Importing Content from other Projects). Otherwise, their coding will be added to yours when estimating statistics such as number of references (e.g., if you both code the same section of text to the same Node, this counts as 2 references).

To create a copy, go to the Share menu tab, click on Copy Project and select the appropriate format in the Copy To area (e.g., NVivo 12 (Windows) project) and click the Browse button to choose where to save it. Then click Save.

The second coder will need to open this copy in NVivo using their own unique User Profile, so that NVivo can identify who coded what. (Note: if the coder uses your computer to code the Project, it will probably open under your User Profile by default. To change User, go to the File menu and select Options. In the User Profile area, change the Name and Initials for the new User. It's also a good idea now to tick the Prompt for user on launch so that you don’t forget to change it back to yourself). Then click Apply and OK.
Coding Comparison Query

You can conduct a **Coding Comparison Query** (on both text and region data types) to get various measurements of inter-Coder reliability, which estimate the **amount of agreement and disagreement between the Coders** for each Node in each document separately (e.g., *Habitat* coding in the *Barbara* interview). These are the notable estimates to check to ascertain reliability:

- **Percentage Agreement** refers to the total number of characters/pixels/timespans in the document on which Coders agree *either to code or not to code*, divided by the total number in the document. If it’s 100%, it’s probably because neither of them coded the document to that Node. If it’s very high (above 80% for example) it could just be an artefact of the large size of the document. **You should always check the following for more clarity on reliability:**
  - **A and B** is the percentage of content coded by both Coders and is a less ambiguous and clearer indicator of actual **agreement** in coding. If this is high (above 70% for example), the coding for the Node in that instance is good.
  - **A and Not B** is the percentage coded by User A only, **B and Not A** is the percentage coded by User B only, and both these together are an unambiguous measure of **disagreement**.

- **Cohen’s Kappa Coefficient** is a statistic which compares the observed level of agreement on what should be coded to that which would have been expected to be coded by chance alone. This ranges from -1 (no agreement) to +1 (perfect agreement). It ignores the amount of the file on which both Coders agreed should not be coded, thus Kappa values can be low even when percentage is high. There are several guidelines for determining what is an acceptable coefficient for reliability. Fleiss et al (2003) suggests <0.4 is poor, 0.41-0.75 is fair to good, and 0.76-1.00 is very good. If Kappa is 1, it’s probably because both of the Coders did not code anything in the document to that Node.

As you will receive **estimates for one Node across multiple documents**, you will want to report the **average or ranges** of percentages/kappa statistics for each Node.

**Example:** conduct a Coding Comparison Query between Effie and Henry for a limited number of Cases and Nodes.

On the **Explore** tab click the **Coding Comparison** icon and select the users you wish to compare in the **Compare coding between** area (e.g., *Effie Reeves and Henry Patterson*). In the **At** box click **Select** and choose the Nodes you wish to include in the analysis (e.g., choose **Nodes in selected sets**, then click the **Select** button and choose the **Nodes for coding comparison** folder). Choose the specific Data Files to use in the **Scope** box (e.g. choose **Data Files in selected sets**, then click the **Select** button and choose the **Interviews for coding comparison** folder). Select the measure(s) you want to use: **Kappa** and/or **Percentage Agreement**. Click **Run**.

You can double-click on a row in the results where there is some disagreement (e.g., Kappa >0 and below <0.3; less than 80%) to compare the coding conducted and see where they disagreed for that File and Node. There will be a coloured coding stripe for Effie’s coding and another colour for Henry’s with their initials attached. The pink coding stripe represents the aggregated coding for that section (Effie and/or Henry’s coding).
Exporting

Exporting Results of Queries
If you want to export the results of Queries, it’s easiest to right-click on the results themselves (e.g., a Word Cloud) and select Export. Choose the file format for the exported file in the Save As Type list. In the File Name box, change the name if you like. Click Browse to select where you want to save it and click OK.

Exporting Nodes
In the List View, select the Node you want to export. On the Share tab, click Export and from the Export Options list select the appropriate option:

- **Entire Content (.html)** will create a webpage with your coded data, which is useful for sharing your project findings (including audio and video clips) with colleagues who don't have NVivo. The Node is exported in two parts: an executable *.html file (e.g., Balance.html) and a folder containing files that make the webpage work (e.g., a Balance folder). When you open the *.html file webpage has a tab for the Node’s Summary, as well as separate tabs for all the coded Text, Audio, Video files etc..

- **Reference View** will export the Reference View to a text document (text only).

- **Summary View** will export the Summary View of the Node to a spreadsheet file.

**Note:** If you want to move a Node HTML file to another location, make sure you copy the folder containing the web page files to the new location as well, or the links will be broken and the webpage won’t work. For Nodes that code audio and video Data Files, the coded sections of media files will be included upon export only if you coded directly on the audio or video timeline.

Click the Browse button under Save As and change the name/location of the exported files if you need. If you choose Entire Content or Reference View there are further options to choose from the Include section of the menu (e.g., include Description, Annotations etc.). Select the Properties and Related Content that you want to include then click OK.

Exporting the List View of the Nodes Area
You can export all the information shown in the List View of the Nodes Area into an Excel file. On the Share tab, click on Export List (or right-click in the List View of the Node area and select Export List). The Excel file includes all the Node Names and the number of Files and References associated with them.

Exporting Classification Sheets
If you want to export a Classification Sheet containing your demographic or bibliographic information for example, you can export it as either an Excel, SPSS or Tab Separated data file. Open on the Classification Sheet you want to export (e.g., the Person classification in the Case Classifications folder). On the Share tab click Export (or right-click on the Classification Sheet and select Export Classification Sheet). Select the appropriate export options: (e.g., export to Microsoft Excel). Choose a location to save the file in the Save As area by clicking on the Browse button and selecting the appropriate place. There are other aspects you can change such as Attribute Value Format (e.g., change Unassigned to (None)) and the formatting of Dates, Times and Numbers. Then click OK.
Other resources

QSR NVivo 12 Help Websites
PC: help-nv.qsrinternational.com/12/win/v12.1.110-d3ea61/Content/welcome.htm
Mac: help-nv.qsrinternational.com/12/mac/v12.1.110-d3ea61/Content/welcome.htm

Training notes
To access training notes, visit the Research & learn webpage anulib.anu.edu.au/research-learn and select the skill area followed by the relevant course. You can register for a workshop and find other information.

Research & learn how-to guides
Explore and learn with the ANU Library’s how to guides (ql.anu.edu.au/howto). Topics covered are:

- Citations & abstracts
- Data Management
- E-books
- EndNote
- Finding books and more
- Finding journal articles and more
- Finding theses
- Increasing your research impact
- NVivo
- ORCID (Open Researcher and Contributor ID)
- Topic analysis
- Using Google scholar from off-campus

Subject guides
Find subject-specific guides (ql.anu.edu.au/subjectguides) and resources on broad range of disciplines. Such as:

- Asia Pacific, Southeast Asia and East Asian studies
- Business, economics, art, music and military studies
- Criminal, human rights and taxation law
- History, indigenous studies, linguistics and philosophy
- Biological, environment, physical & mathematical sciences, engineering & computer science, health & medicine

Navigating the sea of scholarly communication
An open access course designed to build the capabilities researchers need to navigate the scholarly communications and publishing world. Topics covered include finding a best-fit publisher, predatory publishing, data citations, bibliometrics, open access, and online research identity. Five self-paced modules, delivered by international and local experts/librarians (anulib.anu.edu.au/publishing).

Training calendar
A range of workshops are offered to help with your academic research and studies (anulib.anu.edu.au/training-register).