The Word Cruncher

This feature offers word “crunching” capabilities for a simple quantitative content analysis. It creates a list of word frequency counts for the selected or all (currently filtered) textual PDs. A stop and go list and a list of ignorable characters can be used to control the analysis. Further you can create a word cloud for the selected documents, or for all documents:

Creating A Word Count Table

From the main menu, select Analysis / Word Cruncher or click the main toolbar’s Word Cruncher button.

A property dialog opens offering a number of options that affect the procedure:

![Word Cruncher - Settings](image)

Figure 210: Setting properties to start a word frequency count

You may specify a number of properties before starting the frequency count.
Include Selected PD only: Select the scope of the count to be all textual documents or the selected one only.

Built-in Tool: The built-in tool is a simple list window. In addition to the frequency count, size and percentage for each word is shown. The built-in tool can only be used for a selected PD.

Excel: Select this option if you want to export the word count to Excel.

Word Cloud: Select this option to create a word cloud.

Specify content for Excel table: In the next section you can specify the type of content to be included in the Excel table.

Text normalization:

Ignore case: If checked, upper and lowercase letters are ignored: Hello, HELLO, and hello are not counted separately.

Use legacy word recognition: Starting with version 7.0.84, a new way of word recognition has been introduced that improves word recognition for Asian languages. However, the improved recognition pattern slows down the word count. If you were satisfied with the older pattern, click this option to increase speed.

Remove from text before counting: The characters specified in this field will be removed from the source before it is counted. Example: including "eliminates the difference between "Hello" and Hello. Enter special characters that should not be counted.

Use a stop list to exclude words from the count. Instead of enumerating every variant of a word, you can specify patterns. You can create powerful stop lists for different languages, different topics, and different levels of detail.

Deactivate the option “Exclude words” to turn the stop list into a go list. This means, only those words that are in the stop list are counted and all other words are ignored.

The Stop And Go List

Click on the button Edit list if you want to modify it. The first lines starting with a semi colon contain an explanation of terms that can be used. All lines starting with a semi colon are not included in the frequency count.
The terms to be excluded or included can be combined with regular expressions. See below the list of regular expressions that can be used:

<table>
<thead>
<tr>
<th>Expression</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;, &gt;</td>
<td>excludes words consisting of a single character</td>
</tr>
<tr>
<td>/d+</td>
<td>excludes numbers of any length</td>
</tr>
<tr>
<td>-+</td>
<td>excludes strings of hyphens of arbitrary size</td>
</tr>
<tr>
<td>_+</td>
<td>excludes strings of underscores</td>
</tr>
<tr>
<td>0+</td>
<td>exclude sequences of zeros</td>
</tr>
</tbody>
</table>

It is no longer necessary to enter the words to be excluded / included in capital letters. The instruction in the default list still says that all entries must be in capital letters. The reason for this is that we cannot simply exchange the default file as long-time users of ATLAS.ti may already have customized their stop lists. Exchanging the default text file would overwrite all user-defined modifications.

Word Cruncher Output

The figure below shows the result when using the **built-in tool** for a single PD.
If you select the Excel output option, a message pops up informing you that the Excel file can now be opened and where it is stored. Of course, you need to have Excel™ installed to be able to see the result.

Figure 213: Status window information you about word crunch results
Use MS Excel™ functionality like sorting by highest to lowest frequency for further exploration.

Word Cloud

The word cloud output looks as follows:

Fading fades out the less frequent words.

With the slider Limit, you can select how often a word should occur to be displayed in the list.

Right-click on a word and select REMOVE FROM CLOUD ONLY if you no longer want the word to be displayed in the current view.

If you do no longer want the word to be counted in all following runs, select the option: EXCLUDE SELECTED WORD. In stop list mode, this means that the word is added to the stop list. In Go List mode, the word is removed from the Go List.
Switch the sort order by name, weight or word size by clicking on the suitable tab.

The word cloud can also be created by right clicking on a document in the Primary Document Manager. Select **Word Cloud** from the context menu.

## Query Tool

The **Query Tool** is used for retrieving quotations using the codes they were associated with during the process of coding. This is different from a *text* search: To search for occurrences of text that match a specified pattern or string, you have to use the search function or the Object Crawler (see “Text Search” on page 210 and “The Object Crawler” on page 365).

The simplest retrieval of this kind (“search for quotations with codes”) is what you frequently do with the Code Manager: double-clicking on a code retrieves all its quotations. This may already be regarded as a query, although it is a simple one. The Query Tool is more complex in that it can be used to create and process queries that include combinations of codes.

A **query** is a search expression built from operands (codes and code families) and operators (e.g. NOT, AND, OR, etc.) that define the conditions that a quotation must meet to be retrieved (e.g., all quotations coded with both codes A and B).

By selecting codes or code families and operators, a query can be built incrementally which is instantaneously evaluated and displayed as a list of quotations. This incremental building of complex search queries gives you an exploratory approach toward even the most complex queries.

## The Query Tool Window

The Query Tool is launched by clicking the Query Tool button (see left), or by choosing **Analysis / Query Tool** from the HU Editor’s main menu.