10 - Text Search
Contents

10 - TEXT SEARCH ........................................................................................................... 1

SEARCH IN THE FOUR MAIN WINDOWS ........................................................................... 1

LEXICAL SEARCH ............................................................................................................... 2

  Entering a search string ................................................................................................. 3

  Deleting a search string ................................................................................................. 4

  Text search options ......................................................................................................... 4

  Searching with regular expressions ................................................................................ 6

  Saving search terms ......................................................................................................... 7

SEARCH RESULTS ............................................................................................................... 7

  Search results when using the AND combination with search terms ............................. 9

  Opening your last search result ....................................................................................... 9

EXPORT SEARCH RESULTS ................................................................................................. 9

AUTOCODE SEARCH RESULTS ......................................................................................... 12

  Exclude search results from autocoding .......................................................................... 16

EXTENDED LEXICAL SEARCH ........................................................................................... 16

KEYWORD IN CONTEXT (KWIC) ......................................................................................... 18
10 - Text Search

Search in the Four Main Windows

All four main windows in MAXQDA have a magnifying glass icon in their toolbars. Clicking on the **Search** icon opens a search field. Below you can see the search field in the “Document Browser”.

![Search in the “Document Browser”](image)

MAXQDA begins searching as soon as you type the first character into the field and shows you how often that particular search string has been found in the currently opened document. You can use the arrow icons next to the search field to skip through the search results.

By default, all occurrences of the entered text are taken into account, even those that occur in the middle of a word, such as, for example, "less" in "school lessons". The search can be customized using the icons in the search bar:

- **Aa  Case-sensitive** – The letters in the search term and the term found must be the same, including whether they are upper or lower case.

- **a*  Starting letters** - The beginning of a word must match the search characters. The search for "erica" finds the name "Erica", but not "America".

- **abc  Whole word** – The search is checked for word-to-word consistency: if the search term is "teacher", "teachers" is not registered as a hit. This option always includes the option "Starting letters".
Find and replace search terms in texts and tables

If the Edit Mode for a text or table document has been activated, an additional icon appears in the search bar with which you can open a field for replacing the found hits with a new term.

Find & replace – Allows you to enter a term with which you can replace the search hits, for example to render the opened text anonymous.

Please note: As long as Edit Mode is active, the replacements can be undone by clicking the Undo text changes icon in the Edit bar.

Find and replace document and code names

To search within the document and code names listed in the "Document System" and "Code System" and replace them if necessary, proceed as follows:

1. Click on the magnifying glass in the header of the window. The search field will open, in which you can enter a search term. MAXQDA then highlights the hits in the names of your documents and codes.
2. Click on the Find & Replace icon in the header. An additional input field will appear in which you can enter a replacement term.
3. Clicking on Replace or Replace all will replace the selected terms.

Lexical Search

The lexical search function allows you to research your documents without first coding them. You can do searches of either

- documents,
- memos, or
- coded segments, which are currently shown in the “Retrieved Segments” window.
- The search can also be limited to activated documents. You can choose to search in a single document, a document group, or all documents that have certain variable values. The lexical search function can be started by selecting Lexical search from the Analysis drop-down menu or by clicking on the magnifying glass icon in the
“MAXQDA standard” toolbar.

<table>
<thead>
<tr>
<th>Home</th>
<th>Import</th>
<th>Codes</th>
<th>Variables</th>
<th>Analysis</th>
<th>Mixed Methods</th>
<th>Visual Tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reset activations</td>
<td>Lexical Search</td>
<td>Complex Coding Query</td>
<td>Reset Coding Query</td>
<td>Compare Groups</td>
<td>Summary Grid</td>
<td>Summary Tables</td>
</tr>
</tbody>
</table>

Lexical search icon in the Analysis menu tab

After selecting to do a lexical search, the following window will appear:

![The dialog window for the lexical search](image)

**Entering a search string**

When the dialog box first comes up, the left pane is empty and you can just start typing to enter the word or string (called the “search string”) that you want to find. Use the return key to enter multiple search strings.

To add additional search strings later, click "New entry". To edit an existing search string, click on the word and change it.

Search strings may consist of one word, multiple words, or only of a part of a word.
For instance, you can search for “Tom Hanks,” “Hanks,” or only “anks.” Quotation marks are not necessary if a search string consists of more than one word - it would lead MAXQDA to only search for words that are placed in quotation marks.

**Tip:** If you highlight text in the „Document Browser“ then right-click on it and select the option **Search for highlighted text**, the highlighted text will be displayed directly in the Lexical Search window.

Search strings may contain so-called “wildcard“ characters, e.g. * or ?. The characters function as follows:

- If the character “?” appears in a search string, MAXQDA finds, for instance, “Hanks” if the search string is “H?nks,” but “Hunks” or “Hinks” would also be found.
- The character “*” stands for multiple characters, e.g. “H*ks” would find “Hanks” and also “Hawks.”
- The beginning of a word can be delimited by the character “<().” The search string “<in(t)er)” would find “Interest,” “interesting,” but not “winter.”
- The end of a word can be delimited by the character “()>.” Thus, the string “(ks)>” would find “Hanks,” “Banks,” and “thanks,” but not “Bankside.”

To search for the asterisk * or the question mark ?, a backslash “\“ must be placed in front of them: \\* and \\?. In addition, you need to select the option “Interpret search items as regular expressions”.

**Please note:** Hyphenation is not recognized in PDF-documents.

**Deleting a search string**

To remove search strings from the list, click the x symbol next to the word or select the whole line and click the delete symbol at the top of the dialog window. By holding down the **Ctrl key** (Windows) or **cmd key** (Mac), you can delete multiple search strings at once.

**Text search options**

You will find the options for the search procedure in the right pane of the Lexical Search window.

First, you must specify where you wish to search:

- If you select **In documents**, additional selection fields that were previously displayed in light grey will become available. These additional fields include “**Only in activated**
documents” and “Only in retrieved segments”. If no texts are activated, these options will not be available.

Other options:

- **Find whole words** – Search only for strings that are exactly the same as the search string. For example, if you search for “teach,” the string “teacher” will not be included in the search results. If you turn this function off, “teach” would also give you the strings “teacher” or “teaching” as results.
- **Case sensitive** – Search only for strings in which the same letters are capitalized as in the search string.
- **Interpret search items as regular expressions** – the entered search terms are interpreted as “regular expressions”. Regular expressions can be used to perform complex searches (see below for more information).
- **Include word forms from lemma lists** – searches are conducted not only for an entered search term but also for all its word forms. A search for "go" will also find the word "went". The option affects only single words, not composite words. You also have to set the appropriate language for the text.

Please note: You can view and edit the lemma lists used by MAXQDA. In the Help from MAXDictio you can find detailed information about the location of these lists. If you use the word-form function for your search and publish the results, you need to include a license reference in your publication because the used lists are available under an open license.

You can also choose how the search terms are linked with each other:

- With the OR combination, a hit is always included in the results list when one of the search terms is found.
- With the AND combination, all the search terms must occur within a defined distance of each other in an adjustable search range.
- As soon as you have selected AND, another option is made available directly below it:
- **Within xx document/paragraph/sentence.**
  With xx you define the maximum distance of the search terms.
  The following options are available as a search range:
  - **Document** – the two search terms must occur somewhere within a document. In this case, the maximum distance cannot be set.
  - **Paragraph (Text, Tables)** – the two search terms must occur within xx paragraphs. The search only includes text and table documents.
  - **Sentence (Text, Tables, PDFs)** – the two search terms must occur within xx sentences. In addition to text and table documents, the search also includes PDF files.
Examples:

You search for the terms "public" and "citizen" within a document. MAXQDA will then check that both of the two terms occur at least once. A row is displayed in the results table for each document in which both terms occur together.

You search for the terms "public" and "citizen" within 1 paragraph. The list of results then contains as many rows as there are paragraphs in which both search terms occur.

You search for the terms "public" and "citizen" within 2 paragraphs. For example, if paragraph 12 contains the term "public", then "citizen" must appear in paragraphs 11, 12 or 13. Only then will this count as a hit.

As soon as you click the Search button, the search will begin.

Depending on how many documents you search and how complex the search is, it could take some time to find all of the results.

Searching with regular expressions

If you select the option Use regular expressions, MAXQDA interprets the entered search terms as "regular expressions". These are widely used in information technology and allow complex searches to be carried out. The following table contains some examples of regular expression searches:

<table>
<thead>
<tr>
<th>Search term</th>
<th>Explanation</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>\d</td>
<td>Number from 0 to 9</td>
<td>200\d finds all years from 2000 to 2009</td>
</tr>
<tr>
<td>$</td>
<td>Ending letters</td>
<td>cap$ finds all words ending with the letters cap</td>
</tr>
<tr>
<td>[…]</td>
<td>One of the characters in brackets</td>
<td>h[oôô]tel finds hotel, hôtel, hôtel 200[0-9] finds all years from 2000 to 2009 Class [34][a-d] finds class 3a, class 3b to class 4d</td>
</tr>
<tr>
<td>(</td>
<td>)</td>
<td>One of the character combinations in brackets</td>
</tr>
</tbody>
</table>

As an example, a helpful tutorial on the use of regular expressions can be found on the following website: [https://medium.com/factory-mind/regex-tutorial-a-simple-cheatsheet-by-examples-649dc1c3f285](https://medium.com/factory-mind/regex-tutorial-a-simple-cheatsheet-by-examples-649dc1c3f285)
Please note: There are several technical implementations of regular expressions. Perl-compatible expressions (PCRE) are implemented in MAXQDA.

Saving search terms

Searches can be saved, i.e. you won’t have to enter the search terms again. Click Save, enter the name of the saved scan, and then click Save again. Search files are saved with the extension .SEA (an abbreviation for search). Saved searches can be reentered at any time.

Search Results

Your search results will be shown in a table that appears as follows:

![Example search results]

Along with the usual filter functions, the toolbar at the top of the results window provides access to the following functions:

- **Autocode search results** – codes search results with an existing code. All search results that have not been excluded from the selection (and which display a Stop symbol in the first column) will be autocoded.

- **Code search results with a new code** – codes search results with a new code. All search results that have not been excluded from the selection (and which display a Stop symbol in the first column) will be autocoded.

- **Autocode/Export: ignore hit** – exclude selected rows from autocoding and exportation by clicking on this icon.

- **Activated documents with hits** – Adjusts the activation of the documents in the “Document System” so that only those documents are activated for which a search
A term was found. This makes it possible to perform further analyses for these documents, for example, to only consider the activated documents in a group comparison.

- **Detailed list of found locations** - Shows a list containing all individual search hits. This icon is only visible if several search terms were searched for within a specified context (sentence, paragraph, document). For example, if you search for "learning group" and "tutorial" within a given paragraph, the detailed list contains all the hits found individually for "learning group" and "tutorial" where both search terms appear in the same paragraph.

- **Refresh** – Opens the search dialog box again, e.g. to adjust the search settings.

- **Open as Excel table** – the search table will be opened in XLS/X format in the appropriate program (usually Excel). If you have selected a section of the results table (marked green), only these will be included in the new table.

- **Open as HTML table** – the search table will be opened as an HTML table in your standard internet browser. If you have selected a section of the results table (marked green), only these will be included in the new table.

- **Export** – The table will be saved in XLS/X, HTML, RTF or TXT (Tab delimited) format, then opened in Excel or in the default browser. All rows that do not display a stop symbol will be exported.

The table of search results can be handled in the same way as other tables in MAXQDA: clicking on any column header will sort the table according to this criterion. For instance, clicking on **Document** will arrange the document names included in the table in alphabetical order. Clicking here again changes the sequence to descending order. A little triangle indicates the sort attribute and order (ascending or descending).

The entire table or parts of it can be copied to the Windows clipboard (by selecting all or part of it and pressing **Ctrl+C** (Windows) or **cmd+C** (Mac)).

Below the title bar, you can see how many documents contain the search string and how many times the search string appears in those documents.

**Please note:** Once you click on a hit, the corresponding document is opened and positioned exactly where the reference is located. The search term found is highlighted.
Search results when using the AND combination with search terms

If you search for more than one search item within the whole document, a paragraph – or when using the extended lexical search within a sentence –, the search results table will not only list the hits for each search item separately. Instead, it will contain as many entries as combinations of the search items have been found in the defined contexts, thus the number of documents, paragraphs or sentences in which all the search items exist. The following screen shows this for the search for „happiness“ AND „satisfaction“. In the column „Context“ MAXQDA informs you, if these search strings exist within a document, paragraph, or sentence.

A click on a row highlights the first hit of an entered search item within the context.

To get a list of each individual hit of search items, click on the icon Detailed list of hits.

Please note: When exporting the search hits or when autocoding them, MAXQDA uses the detailed list of search items.

Opening your last search result

As long as your project is still open, you can always go back and review the last search result, thus saving time in case you need to do the same search again. Go to the Analysis menu tab, click at the bottom of the Lexical Search entry to open a short drop-down menu and select Last Search Result.

Export Search Results

The results table of the lexical search tool offers the possibility to export search results by clicking on the Export icon in the toolbar of the results window. The following
options window will appear, where the range of the data to be exported in relation to the search term can be defined:

PDF documents
- Only search string
- Sentence
- Paragraph

Context
- Words before
- Words after

Text and table documents
- Only search string
- Sentence
- Paragraph

The export window is divided into two parts. Options for the export of search results from a PDF document can be selected in the upper section, with the corresponding options for search results from text and table documents in the lower section. Exports of pure search results as a simple list is generally not useful; therefore you can define the range for the export of search results.
For PDF documents, you can define how many words or sentences before and after the actual search term are exported. This option is also available for text and spreadsheet documents, in addition to the **Paragraph** option. When you select one of these two options, the name of the context options will change and you can choose how many sentences or paragraphs before and after the search term you want to include:

<table>
<thead>
<tr>
<th>Text and table documents</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Context Options" /></td>
</tr>
</tbody>
</table>

Setting **Paragraph** with **0 Paragraphs before** and **0 Paragraphs after** means that only paragraphs in which the search term is found will be exported. If the term appears several times in the same paragraph, the paragraph is still listed only once, whereas the search term or terms will be listed more than once.

Setting **1 Paragraphs after** means that the paragraph containing the search term, as well as the following paragraph, will be exported.

**Please note:** Read the chapter on **Extended Lexical Search** to learn how MAXQDA recognizes sentences.

When you click **OK**, a window will pop up in which you can select the file name, format, and location to be saved. The following options are available:

- Excel (XLS/X)
- Website (HTML)
- Maintext in Rich Text Format (RTF)
The result as Maintext in Rich Text Format will appear as follows:

Search string: family
to go to college straight from high school in 1964 and become a teacher. I was a
good student in high school graduating 19th in a class of 403. I really wanted to go
to college, but my family couldn’t afford it. Unfortunately, my mother had the
attitude college was a waste for girls, because they got married and didn’t put their
college education to use. So, I never pursued the issue. I know, I shouldn’t blame my
mother for my not going to college. I could have been more enterprising and tried to
find a way to pay for it on my own. I think never having been a teacher will always
haunt me. Now it’s too late. I’d have to take a large salary cut if I were to change
professions at this stage in life.

Each search result consists of the source data, including the name of the document
group and document, and the position in which the search string was found (in
brackets), followed by the search string itself.

Please note: The recognition of sentences in PDF documents is optimized for left-to-
right languages and may not work as expected with right-to-left languages.

Autocode Search Results

Like the segments in the “Retrieved Segments“ windows, lexical search results can be
automatically coded, meaning all text passages found in the search are assigned as
specific code. This automatic coding has the advantage of being easy and reliable. The
disadvantage, however, in comparison to manual coding, is of course that you can’t
examine the relevance of each respective text passage during the coding process. The
occurrence of a specified string in the text automatically triggers the coding. For
example, if the word “Mother“ is automatically coded with the code “Family“, references to “Mother Earth“ would also be included.

To code the search results automatically, follow these steps:

- The best way to code search results automatically is to click on the Autocode search
  results with a new code symbol.
- In the dialog box that appears, enter the new code and click OK. The new code will
  be inserted at the top of code system.
In the ensuing options window you can set the range before and after the search term as well as the weight, which is established during the coding process.
In the "Code" field, you establish with which code MAXQDA should assign to the search results. The “Quick list of codes” – which lists the most recently used codes – is
displayed in this window. The code that you created earlier is automatically selected here. Furthermore, a weight value can be set for all coded elements, that is to say, all segments will be coded with the same weight.

In the “PDF documents” field, you can set whether only the respective search result should be coded, or whether additional words or sentences before and after the search results should be coded as well.

Please note: Read the chapter on Extended Lexical Search to learn how MAXQDA recognizes sentences.

In the “Text and table documents” field, you can also determine if whole sentences or paragraphs before or after the search result should be coded.

Setting the parameters for Autocoding

Setting Paragraph with 0 Paragraphs before and 0 Paragraphs after means that only the paragraph in which the search term is found will be coded. If the term appears several times in one paragraph, it will nevertheless be coded only once.

The setting 1 Paragraph after means that the paragraph in which the search term is found as well as the following paragraph will be coded. Again, if the term appears several times in one paragraph, it will nevertheless be coded only once.

Please note: If you do not want to code with a new code, but with an existing code from the “Code System”, you can use the Autocode search results option. Before you start the process, click on the desired code in the Code System, which will cause the code to appear in the list of recently used codes in the Autocode options window.
Exclude search results from autocoding

Fundamentally, all search terms that appear in the results list will be autocoded. However, before executing the autocode function, you can review the list of search terms in order to decide whether each term should be included or not. Once one or more lines in the results table have been selected, click on the red icon _Autocode/Export: ignore hit_ in order to exclude the entry.

The fastest way to exclude a term is to double-click in the first column.

Please note: The selected codes are incorporated into the undo coding list which can be accessed from the Code toolbar with the ⦿ icon. This action can be reversed with a single command if necessary.

Extended Lexical Search

The _Analysis > Extended lexical search_ function, as with the simple search option, allows you to analyze your documents to a certain extent without having to have already done any coding. In addition to the simple search, the extended search offers to search for different combinations of search words. The searches can be carried out
in:

- documents,
- memos, or
- the coded segments currently called up in the “Retrieved Segments” window.

The search can also be limited to those documents that are currently activated in the “Document Browser”.

To start an extended lexical search:

1. Go to the **Analysis** menu tab and click at the bottom of the Lexical Search entry.
2. From the drop down menu select **Extended Lexical Search**.

![Extended lexical search dialog box](image)

The dialog is separated into several parts. At first you can enter search items in the text fields.
• All these search strings (AND combination)
• Any of these search strings (OR combination)
• None of these search strings (search strings for exclusion)

Please note, that you can separate different search strings with a space. If you want to enter a combined search string you can use quotation marks like “Mahatma Gandhi” around your search string. Question mark (“?”) and asterisk (“*”) work as placeholders for a single character or for one or more charates. Use <(word) for word begin and (word)> for word end, <(inter) will find “interest” and “interesting”, but not “winter”, and (ks)> will find “banks” and “thanks”, but not “bankside”.

**Tip:** Please be aware that hyphenation can not be recognized in PDF-documents.

In the upper right part of the dialog you can define, in which part of the project MAXQDA will search an in the options on the right you can define in which context the search items have to exist.

All three search entries will be combined with AND. In the screen above MAXQDA will present all sentences in which “climate” and “energy” exist together, but not “politics”.

**Please note:** Sentences are defined by MAXQDA according to the following rules: A sentence always begins following a period, question mark, exclamation mark, or colon. The following exceptions apply:
# A number that is not four digits appears before a period (e.g. 1. or 2.).
# A single character appears before a period (to exclude abbreviations).
# Two identical characters appear directly before a period (e.g. ff. or pp.).
# Literal speech in quotation marks belonging to the sentence itself.
# First letter after a sentence is in small letters.
# A number appears directly following the end of a sentence.
# Quotation marks appear immediately after the end of a sentence.
In text or table documents, a new sentence invariably begins after a paragraph.

The recognition of sentences in PDF documents is optimized for left-to-right languages and may not work as expected with right-to-left languages.

**Keyword in Context (KWIC)**

Often, you wish to determine the context in which a given term is used. In this case, you can create and print a so-called KWIC (KeyWord In Context) listing. The procedure is as follows:
1. Perform a lexical search by keyword.
2. Double click on any search result in the result table to exclude this result (a stop sign symbol indicates which lines are excluded).
3. You now have two option: Export all KWIC-listing to XLS/S, HTML, or RTF by clicking the **Export** icon. In a second step you can print this file. Or click the **Autocode search results** icon to code the search results with a new code, e.g. "KWIC". You could then double-click the KWIC code to see the overview of coded segments and continue your analysis from there.

In both cases, a dialog window will appear and let you decide how much context to include, e.g. 1 sentence or 10 words or after the search string.

The following image shows what the export of a keyword-in-context search looks like. It has the following structure: each hit starts with the source data (document group, document name, and the paragraph number containing the hit). The next line indicates the search word or search string that has been found in that paragraph. Finally the text segment itself is listed.
I also wanted a job that I would be free to take time off to have a family and not work at the same time, I know that this kind of job does offer that and that makes it easier to be happy with this career choice. I wanted to get a "good" job because I go to college for a reason, but I don't want a job that is better paying than my husband because I believe that the man in a family is the one to be working and not the woman.

We can talk to each other about everything. I have been welcomed into her family as another son or brother.