

Making Use of Word-based Features in your research

Michael C Gizzi (mgizzi@gmail.com mgizzi@ilstu.edu)

This session explores how you can take advantage of word-based features in your work, focusing on **Word Explorer**, a tool that aids in identifying concepts, auto-coding, refining codes and more throughout the research process.

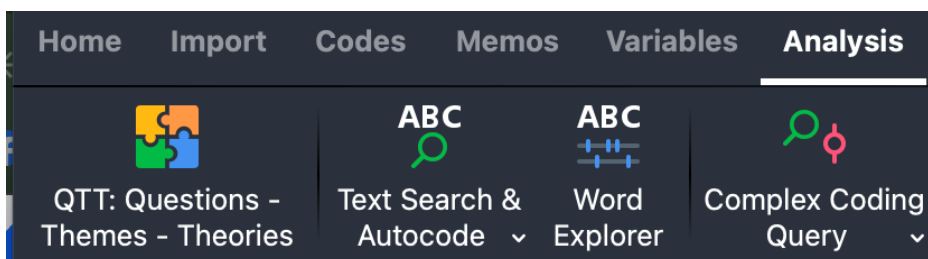
Word-based features cover a lot of territory in MAXQDA, from text searches to MAXDictio functions such as word frequencies, word combinations, keywords in context, word matrix browser, and dictionary-based content analysis. It brings together the powers of a text search function with all of these options in one place, providing a one-stop tool to see the way words and phrases appear in your data, and to easily find segments and auto-code search results.

What does Word Explorer do?

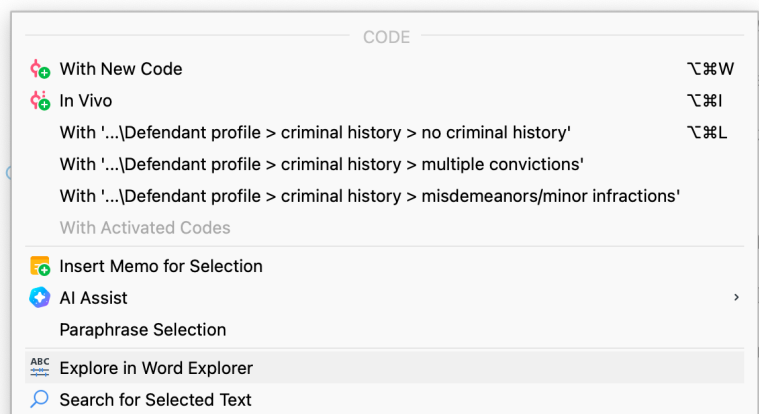
- It lets you contextualize the usage of specific words or phrases in documents, codes, code comments, memos, and paraphrases
- Identify the context words and phrases are used in combination with
- Select any element to see the actual underlying data, and use auto-code tools

How to access Word Explorer.

Go to Analysis > Word Explorer

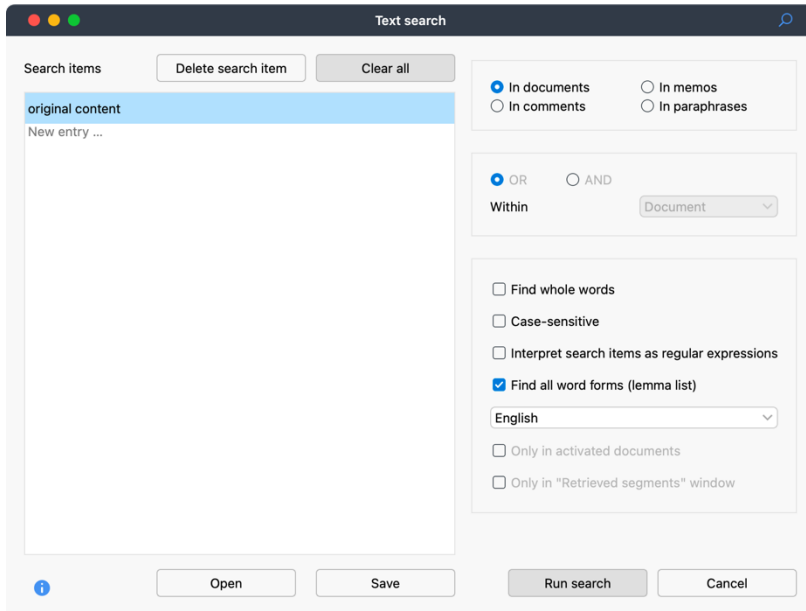


Or highlight a segment of text in a document, right click and select “Explore in Word Explorer”



Contrasting Text Search and Word Explorer

The classic tool to search words in MAXQDA is **Text Search** (formerly known as Lexical search in prior versions). Both tools are in the Analysis menu. As an example, let's take the phrase "original content" and search it first in Text Search & Auto-code.



You select words or phrases, and then choose options such as "in documents, comments, memos, paraphrases"

You can limit the results to whole words, case-sensitivity, base form (lemmatization), and activated documents/codes.

The search creates a search results window, that shows each segment of text found, and provides the ability to auto-code segments.

MAXDAYS

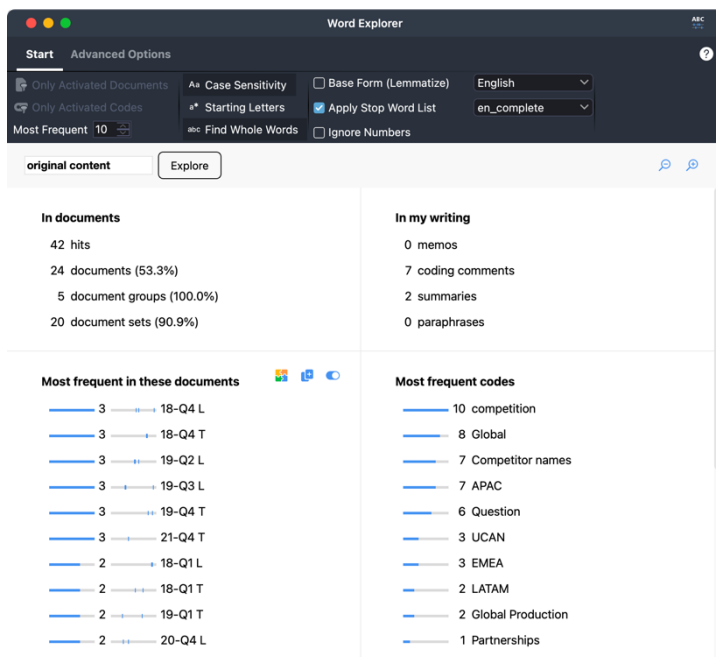
MAXQDA VIRTUAL CONFERENCE

Search results

ANY: original content 45 hits in 24 documents and 5 document groups

Preview	Document group	Document name	Search item	Beginni...	End
We continue to forecast free cash flow of -\$3 to -\$4 billion in 2018, and to be free cash flow negative for several more years as our original content spend rapidly grows.	2018	18-Q1 L	original co...	25	25
We have about \$2.6 billion in cash and we will continue to raise debt as needed to fund our increase in original content.	2018	18-Q1 L	original co...	26	26
the world, it's mostly about picking great people, giving them a great place to work, trusting their -- trusting them and empowering them to continue to make great choices, which we've really been focused on, both producing original content in the U.	2018	18-Q1 T	original co...	99	99
so successful for us all around the world that we've bought the series IP, and we'll be producing sequels and spinoffs of Casa De Papel as original content for years to come.	2018	18-Q1 T	original co...	125	125
In the short to medium term, we're still licensing content off net from them, and they're also producing original content for us, like Nurse Ratched from Fox or the Marvel series from Disnev.	2018	18-Q2 T	original co...	46	46

Text Search is quite powerful and provides numerous ways to search, but when contrasted with **Word Explorer**, we can see how the latter will search words and



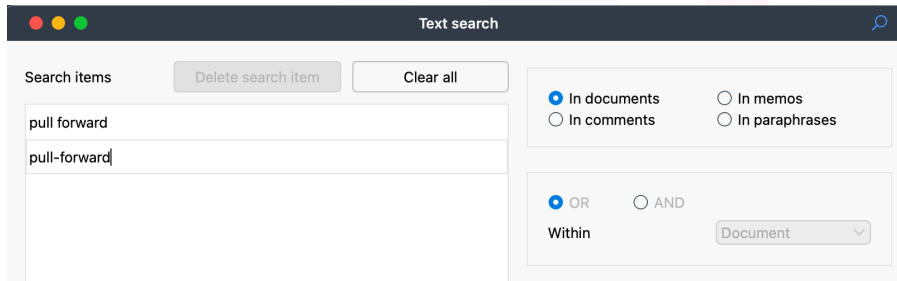
phrases in a way that goes beyond a simple text search.

The default use of the tool provides a list of all results, in terms of hits, documents, document groups, sets, as well as where the search term is found in memos, coding comments, summaries, and paraphrases.

It also shows most frequent documents, codes, word combinations, and context words.

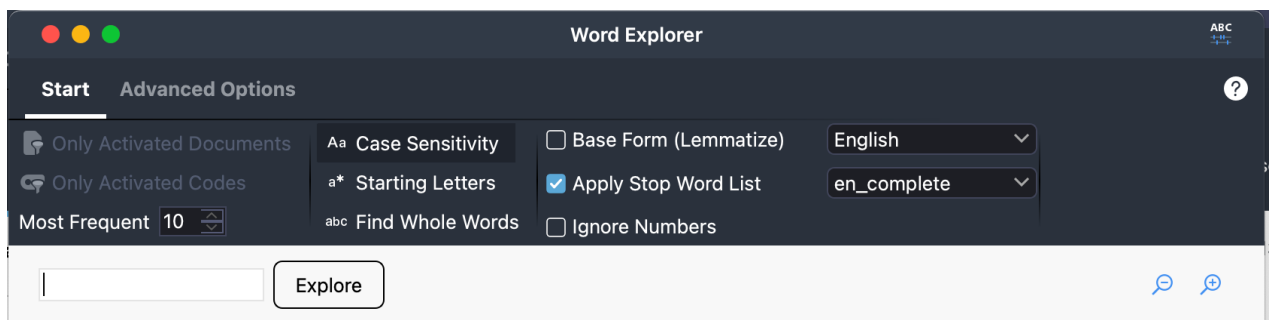
From there you can see the actual search results, activate documents, or view coded segments.

The one thing Word Explorer can't do that Text Search & Autocode can, is to a complex search with multiple search terms. For example, if you have data in which a phrase like "pull forward" is also written as "pull-forward," Text Search could have two separate search items to collect both results.



This will provide all of the segments meeting both versions of the term.

Word Explorer – A look at features

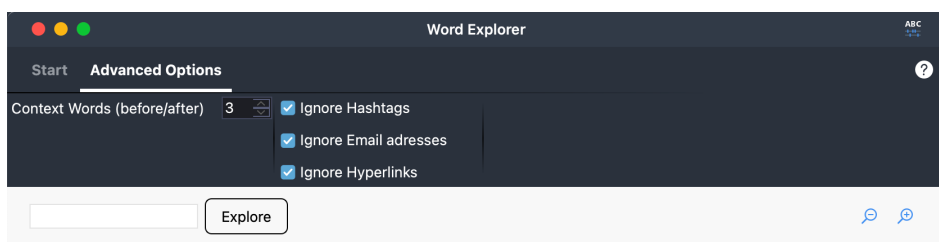


Options:

1. **Only activated documents or codes** Narrow the search to activated documents or codes. This is useful if you want to only search a subset of your data
2. **Most Frequent.** Choose how many results you want to show up for results in the categories of most frequent documents, codes, word combinations, and context words. Range is from 3 to 50.
3. **Case Sensitivity.** Choose if you want the results regardless of capitalization.
4. **Starting Letters.** This is useful if you want to get more than literal word results. For example, if you use the word “bundle” as a whole word, you would not get bundles or bundling. If you search for “bundl” with Starting Letters, the results will include all words that start with those words.

- 5. Find whole words.** This is the inverse of Starting letters. It will produce results for just the exact word you are searching on. If Find whole words and Case Sensitivity is on, it will further narrow the results.
- 6. Base Form.** Also known as lemmatization, base form is a very useful option that will count words from the same stem together. For example, if you search for the word “give,” Word Explorer will include occurrences of give, gave, or given. Base form is currently available for 14 languages. Make sure you select the proper language with the drop-down box.
- 7. Stop List.** Selecting this option will use either a generic or customized stop list. A set of uninteresting words, for example, words such as to or and, or the, which you wish to exclude from the Word Explorer results. When you check this option, you need to specify which stop list to use. The stop list can be edited in MAXDictio > Stop Lists. There is an “english_complete” stop list available from the MAXQDA Help program for the feature.
- 8. Ignore numbers.** This will only return text and not numbers in word combinations containing a number and context words consisting only of a number will not be output. But numbers entered in the search “explore” field are never ignored.

Advanced Options



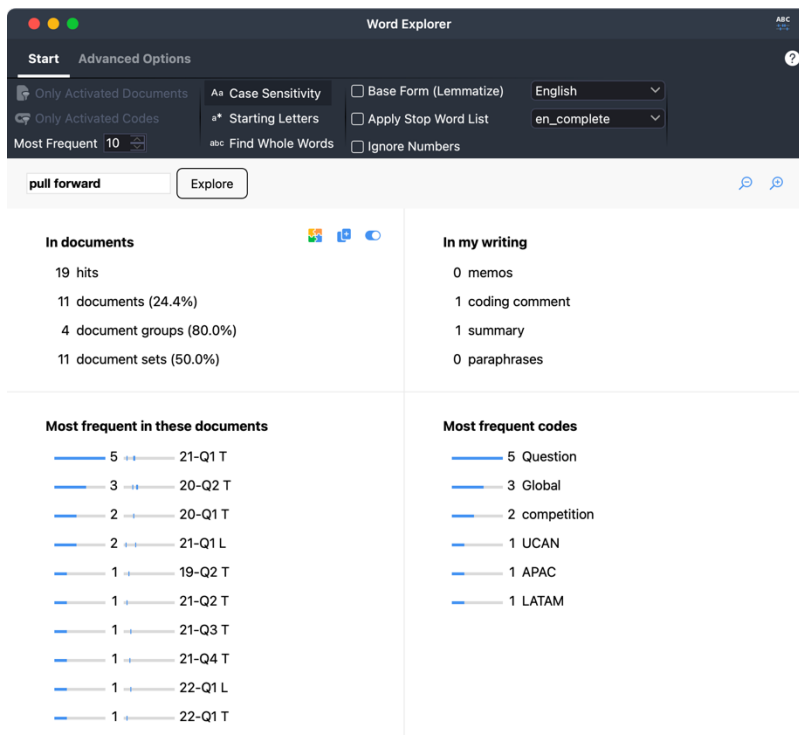
There is a second tab of advanced options. These include:

- 1. Context Words (before/after) [number 1-15]** – here you can set the maximum number of context words that will be considered before and after the search hit. By default, 3 words before and after search hit are considered as long as they belong to the same sentence.
- 2. Ignore hashtags, email addresses, and hyperlinks.** These check boxes let you

choose to ignore hash tags, email addresses, and hyperlinks from results.

Using Word Explorer

Lets take the example “pull forward” and process it through Word Explorer, selecting Select whole words. We do not need base form or stop lists enabled.



This tells us that there are 19 hits for the term, in 11 documents. The phrase appears in one coding comment and 1 summary.

It then also tells us the documents where the phrase occurs most frequently, as well as the most frequent coded segments the phrase occurs within.

If you scroll down in the word explorer window, you can also see the most frequent word combinations and context words. Compared with a Text Search, Word Explorer is providing us with a comprehensive look at where our search terms occurs and its context. It draws on the tools of MAXDictio as well.

Most frequent word combinations

- 4 pull forward in
- 3 pull forward in 2020
- 3 pull forward that you
- 3 pull forward of
- 3 pull forward that
- 3 the pull forward
- 2 COVID pull forward
- 2 of pull forward
- 2 that pull forward
- 1 big Covid-19 pull forward

Most frequent context words

- 13 of
- 9 the
- 6 that
- 5 in
- 3 you
- 2 a
- 2 and
- 2 big
- 2 COVID
- 2 our

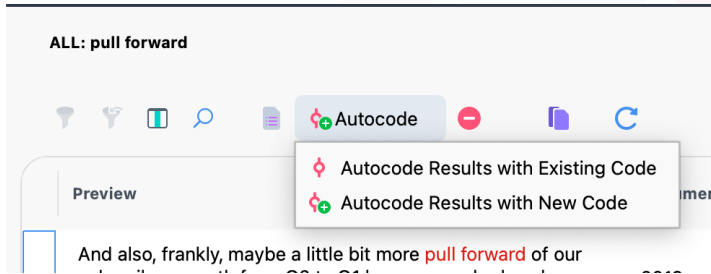
Let's click on the "In Documents" result of 19 hits. This provides us with a search result.

Search results

ALL: pull forward 19 hits in 11 documents and 4 document groups

Preview	Document group	Document name	Search item	Beginni...	End
And also, frankly, maybe a little bit more pull forward of our subscriber growth from Q2 to Q1 because we had such a strong Q1 with 9.7 million paid net adds.	2019	19-Q2 T	pull forward	25	25
It's essentially a pull forward of the rest of the year.	2020	20-Q1 T	pull forward	36	36
So we have had a number of questions about the member growth and the topic of pull forward that you just referenced, Reed.	2020	20-Q1 T	pull forward	38	38
Obviously, we've seen some pull forward in our member growth, if you will, but frankly, our strategy hasn't fundamentally changed.	2020	20-Q2 T	pull forward	43	43
And so when you think about your guidance in that context, it does seem to embed an expectation of a lot of pull forward of growth.	2020	20-Q2 T	pull forward	50	50
There's just that kind of near-term pull forward that you're seeing.	2020	20-Q2 T	pull forward	53	53

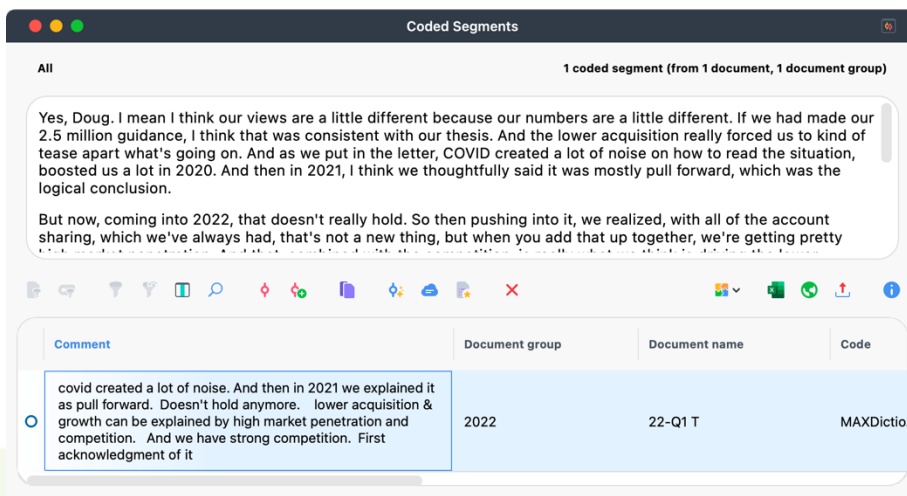
If we wanted to we could easily autocode that search result with a new code for "pull forward" by clicking on autocode. This would let us choose to code with an existing code or create a new one.



We can also eliminate search results that are either irrelevant or ones we do not want to include by single clicking the “=” icon in the left most column to exclude it. A stop icon will occur. Or we could highlight individual rows and use the same ignore icon from the menu.

	Preview	Document group
=	And also, frankly, maybe a little bit more pull forward of our subscriber growth from Q2 to Q1 because we had such a strong Q1 with 9.7 million paid net adds.	2019
⊘	It's essentially a pull forward of the rest of the year.	2020
=	So we have had a number of questions about the member growth and the topic of pull forward that you just referenced, Reed.	2020

Finally, we can use the activate documents containing the search results icon if we want to just use those for analysis.



Looking at results “in my writing” will not produce search results, but instead the coded segments that match the term. For example, this image shows us the segment and the comment with the word pull forward.

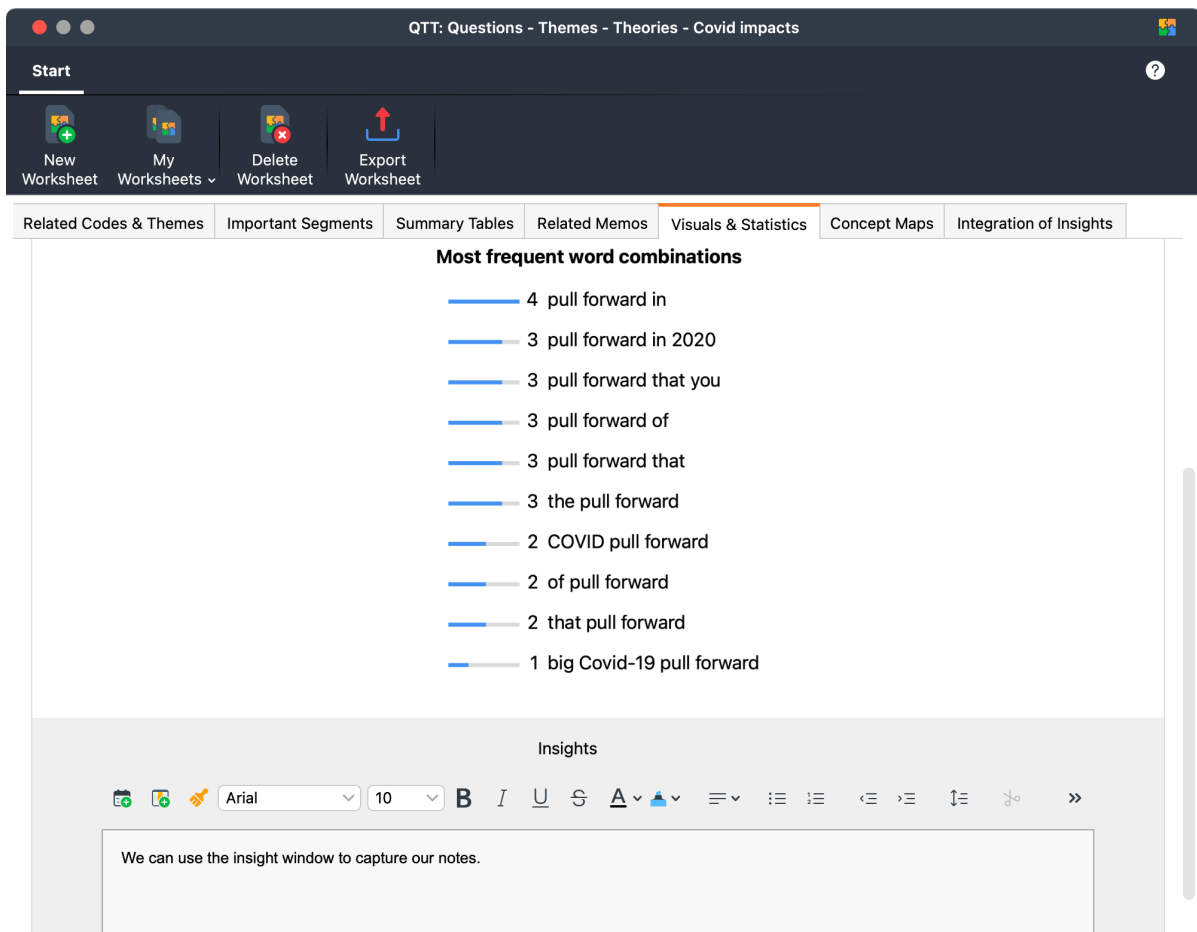
Most frequent word combinations



- 4 pull forward in
- 3 pull forward in 2020
- 3 pull forward that you
- 3 pull forward of
- 3 pull forward that
- 3 the pull forward
- 2 COVID pull forward
- 2 of pull forward
- 2 that pull forward
- 1 big Covid-19 pull forward

When we click on any of the Word Explorer windows, we will also see options in the right hand corner to save the results to a QTT worksheet, copy and image of the results, or disable that particular tool.

If we select the Send to QTT worksheet, we can specify what worksheet to use, and it will transfer the information that worksheet. You can open it and add insights or notes from Analysis > QTT



Please note, though, in QTT, the results are just an image, and not interactive like they are in the Word Explorer window.

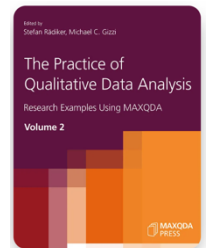
Conclusion

Word Explorer is a powerful tool that brings together several MAXQDA features and can be used as a way to explore your data, search for specific terms, and even auto-code data.

It is a great MAXQDA feature to use if you are developing a dictionary of terms to use in a MAXDictio Dictionary-based content analysis. It does have a few limits, such as the ability to only search on one set of words or phrases and not do complex boolean lexical searches. But it offers the user so much more by letting you see the whole picture of how a word or phrase appears in your project.

Word explorer is also a great feature to use when you identify a new code half-way through your coding process. Using a series of word explorer searches can help you go back to prior documents (activate those first) and look for relevant passages that would fit the new code.

For more examples of how to use Word Explorer, I'd recommend you download a free copy of The Practice of Qualitative Data Analysis With MAXQDA: Research Examples Using MAXQDA, Volume 2 from MAXQDA-Press. Then read chapter 3, "Using MAXQDA for Dictionary-Based Content Analysis: An Example Exploring the Business Strategy of Netflix Over Time." In this chapter word explorer is described and used several times.



For More information contact. Dr. Michael Gizzi,
Certified MAXQDA Trainer

<https://www.4stagesofresearch.com> mgizzi@ilstu.edu or mgizzi@gmail.com