

# Generative AI for Qualitative Analysis

**CARMA Online Short Course**  
**2<sup>nd</sup> – 5<sup>th</sup> June 2025**

**Christina Silver, PhD. FAcSS. SHFEA**

Associate Professor (Teaching), Dept. Sociology, University of Surrey  
Director of the CAQDAS Networking Project  
Director of QDA Services

**Linktree:** [https://linktr.ee/Christina\\_QDAS](https://linktr.ee/Christina_QDAS)

**Bluesky:** @christinaqdas.bsky.social

**Mastodon:** @ChristinaSilver@sciences.social

**Twitter:** @Christina\_QDAS

**LinkedIn:** [linkedin.com/in/christina-qdas/](https://www.linkedin.com/in/christina-qdas/)

**Instagram:** @christina\_qdas

**Facebook:** [facebook.com/QDAServices](https://www.facebook.com/QDAServices)

**YouTube:** [youtube.com/@ChristinaSilverPhD](https://www.youtube.com/@ChristinaSilverPhD)

**Podcast:** [anchor.fm/christina-qdas](https://anchor.fm/christina-qdas)

**Website:** [qdaservices.co.uk](https://qdaservices.co.uk)

**Email:** [christina@qdas.co.uk](mailto:christina@qdas.co.uk)



1



## Day 3

Using AI for qualitative data analysis

2

# GenAI for Qualitative Data Analysis

Methodologically appropriate use of AI for QDA

GenAI use throughout the qualitative research workflow

AI summarization

AI coding

AI conversing

3

## GenAI throughout the Qualitative Research Workflow

**AI tools for qualitative analysis**

- integrations into existing tools (e.g. ATLAS.ti, MAXQDA, NVivo, QualCoder, WordStat, QDA Miner)
- new Aps (e.g. AILYZE, Reveal Qinsights, Tailwind)

Research Design

Literature Review

Data Collection

Transcription

Familiarise

Categorise

Analysis

Interpret

Write up

4

# GenAI tools for qualitative analysis

Summarisation

Categorisation

Conversing

5

Note; detail subject to change as tools develop

| Established CAQDAS-packages         |                                    |                                 |                                    |  |   | New Online Applications  |  |
|-------------------------------------|------------------------------------|---------------------------------|------------------------------------|--|---|--|--|
| ATLAS.ti<br>1989/2023               | MAXQDA<br>1989/2023                | QualCoder<br>2019/2024          | NVivo<br>1987/2024                 | WordStat<br>1998/2025  | QDA Miner<br>2004/2025  | AILYZE 2023<br>CoLoop 2023<br>Reveal 2024<br>QInsights 2024<br>Tailwind 2025 |  |
| Transcription                       | Transcription                      |                                 | Transcription                      | Text Extraction<br>Summarisation<br>Topic Extraction<br>Topic Naming<br>Topic Grouping<br>Syntactic phrase classification<br>Named Entity classification<br>Custom scripts | Text Extraction<br>Summarisation<br>Code Similarity<br>Searching<br>Table Filtering<br>Follow-up AI queries | (Transcription)  |  |
| Summarisation<br>(multiple levels)  | Summarisation<br>(multiple levels) |                                 | Summarisation<br>(multiple levels) |  |   | Summarisation<br>(transcript level)  |  |
|                                     | Explanations                       |                                 | Explanations                       |  |   |  |  |
| Code Suggestions                    | Code Suggestions                   |                                 | Code Suggestions                   |  |   |  |  |
| Coding<br>(automatic & intentional) | Coding<br>(1 doc/code at a time)   | Coding<br>(search & code-based) |                                    |  |   |  |  |
| Chat<br>(conversational)            | Chat<br>(conversational)           | Chat<br>(code/topic/ general)   |                                    |  |   | Chat<br>(conversational & grid)  |  |

and lots of others

more may appear

NO CODING

Plus automatic analyses in some new APs, e.g. MyRA

6

Implementation of AI summarization

Methodological considerations:

- Automated vs user-initiated summarization
- Summarization level
- Summarization length
- Input/output language
- Ability to guide summarization via prompting

- what are summaries based on?
- when might they usefully be used?
- what do they contribute to the analytic process?
- are they linked to the data they are derived from?

9

When, why and how to use AI summaries in the QDA workflow?

Methodological considerations

- what are summaries based on?
- when might they usefully be used?
- what do they contribute to the analytic process?
- are they linked to the data they're derived from?

- The 'raw' data or the work you've done with the data (i.e. via codes)
- Either way, they're based on the form of the text, NOT the meaning

At the beginning of a project – before any human analysis has taken place?

During analysis – as part of the process?

Later on – as a way of synthesizing?

What has the role of summarization been in QDA previous to Gen-AI? Are we doing this just because we can?

If not, is this problematic from ethical and methodological perspectives?

11

CARMA Online Short Courses  
www.qdas.co.uk

Christina Silver, PhD  
christina@qdas.co.uk



14

## Coding qualitative data

### What is coding?

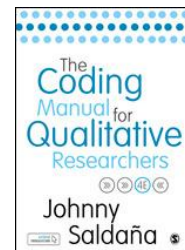
“The process by which segments of data are identified as relating to, or being an example of, a more general idea, instance, theme, or category”

(Silver & Lewins 2014:158)

Thinking about and using...  
be an essential companion  
researcher. Using a number  
methods. Software, providing  
practical in-depth research  
Qualitative Research, provides  
a step-by-step guide.  
Assisted Qualitative Data  
(CAQDAS). Christina Silver  
& Ann Lewins. An essential  
practice and principles of  
to the second edition.

### What is a code?

- “...a word or short phrase that symbolically assigns a summative, salient, essence-capturing, and/or evocative attribute for a portion of language-based or visual data” (Saldana 2016:4)



15

# Coding methods (Saldaña 2021)

### Grammatical Coding Methods

- Attribute coding
- Magnitude coding

### Elemental Coding Methods

- Structural coding
- Descriptive coding
- Process coding
- In-vivo coding
- Initial coding
- Concept coding

### Affective Coding Methods

- Emotions Coding
- Values Coding
- Versus Coding
- Evaluation coding

### Literary & language Coding Methods

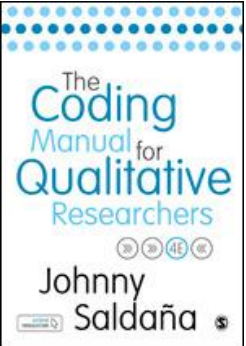
- Holistic coding
- Dramaturgical coding
- Motif coding
- Narrative coding
- Metaphor coding

### Exploratory Coding Methods

- Provisional coding
- Hypothesis coding
- Eclectic coding

### Procedural Coding Methods

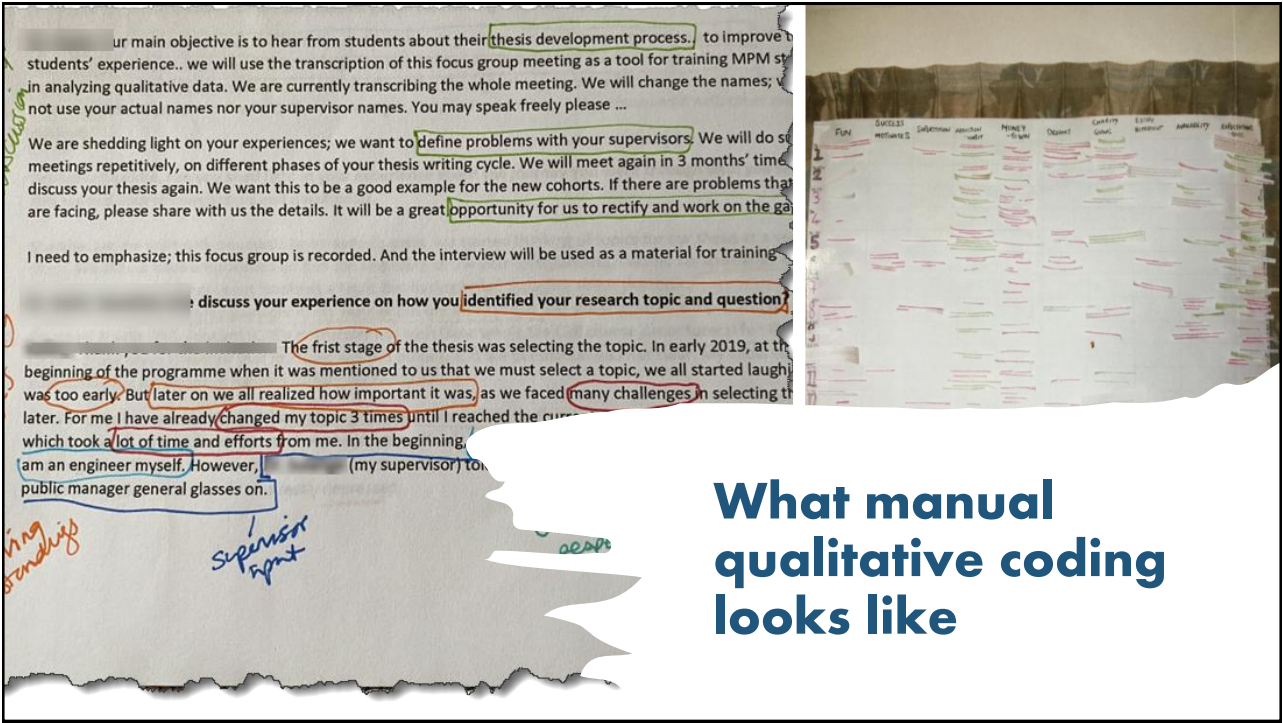
- Protocol coding
- Domain & Taxonomic coding
- Causation coding



### Methods of theming the data

- Categorically
- Phenomenologically
- Meta-summary & meta-synthesis

16



ur main objective is to hear from students about their thesis development process, to improve students' experience.. we will use the transcription of this focus group meeting as a tool for training MPM students in analyzing qualitative data. We are currently transcribing the whole meeting. We will change the names; we will not use your actual names nor your supervisor names. You may speak freely please ...

We are shedding light on your experiences; we want to define problems with your supervisors. We will do so in meetings repetitively, on different phases of your thesis writing cycle. We will meet again in 3 months' time to discuss your thesis again. We want this to be a good example for the new cohorts. If there are problems that you are facing, please share with us the details. It will be a great opportunity for us to rectify and work on the gaps.

I need to emphasize; this focus group is recorded. And the interview will be used as a material for training.

discuss your experience on how you identified your research topic and question?

The first stage of the thesis was selecting the topic. In early 2019, at the beginning of the programme when it was mentioned to us that we must select a topic, we all started laughing because it was too early. But later on we all realized how important it was, as we faced many challenges in selecting a topic. For me I have already changed my topic 3 times until I reached the current topic which took a lot of time and efforts from me. In the beginning, I am an engineer myself. However, (my supervisor) told me to wear public manager general glasses on.

*ing standing*  
*supervisor point*  
*depts*

## What manual qualitative coding looks like

17

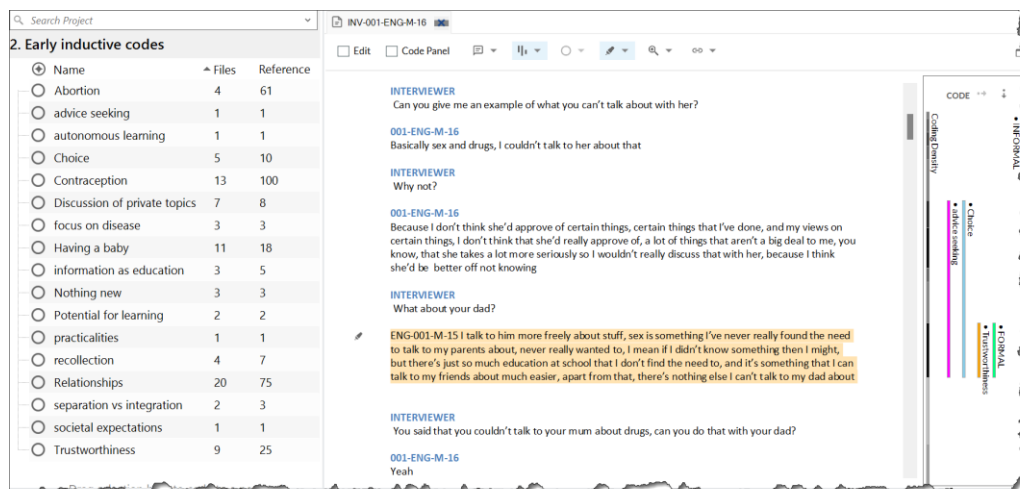
CARMA Online Short Courses

www.qdas.co.uk

Christina Silver, PhD

christina@qdas.co.uk

## What coding looks like in CAQDAS (NVivo example)



18

## Implementations of AI Coding

**Traditional-AI tools** (session 3)  
can result in coding  
(sentiment analysis, pattern-based  
auto-coding, text classification)

In addition, Automated Theme  
Coding (coming up)

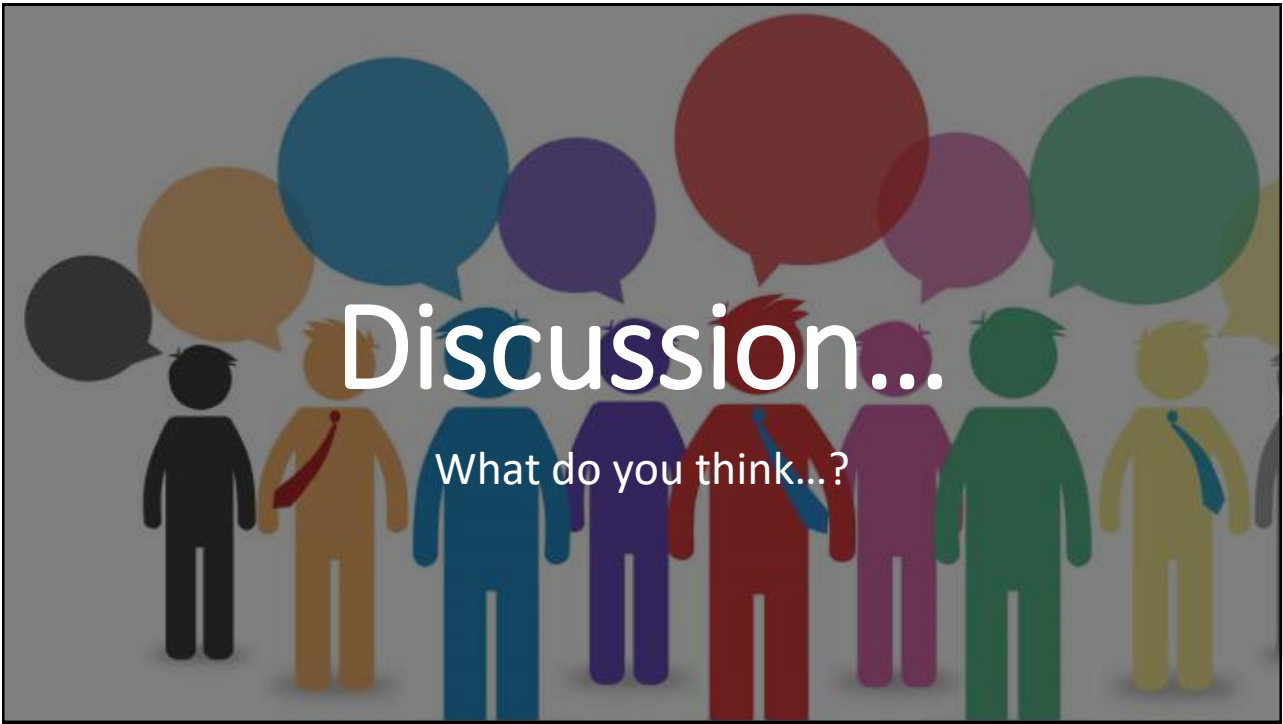
### Gen-AI coding tools

- Code suggestions
- AI Coding
- AI Intentional Coding

### Methodological considerations

- How does AI coding compare to human coding?
  - Speed / interpretive value (is it *really* "coding"?) / number of codes / defining codes
- How is AI coding integrated with human interpretation?
  - input to the process / sequencing of input
- Do we still need to code considering other (new) options?

19



20

# When, why and how to use AI coding in the QDA workflow?

## Methodological considerations

- what is the coding based on?
- when might it usefully be used?
- what does it contribute to the analytic process?
- is it linked to the data they're derived from?

- The 'raw' data or the work you've done with the data (i.e. via already coded materials?)
- How is it determining what a code is? Does this match up with your methodological needs?

At the beginning of a project – as a form of familiarisation?  
Instead of human coding - as the process?  
In combination with human coding – to generate additional ideas?

What has the role of coding been in QDA previous to Gen-AI? Are we doing this just because we can?

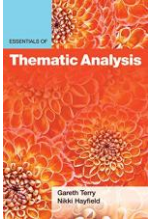
If not, is this problematic from ethical and methodological perspectives?

21

# What are themes?



- Patterns of meaning across a dataset that address a research question (Braun & Clarke 2021)
- Thematic Analysis = popular set of methods



24

# Comparing analysis methods

## Reflexive Thematic Analysis (Braun & Clarke 2021)

1. Familiarise
2. Coding
3. Generate initial themes
4. Develop & review themes
5. Refine, define & name themes
6. Write up

## Template Analysis (Crabtree & Miller 1992)

1. Familiarise
2. Data Coding
3. Formulate coding template
4. Apply template
5. Adapt template to data
6. The final template

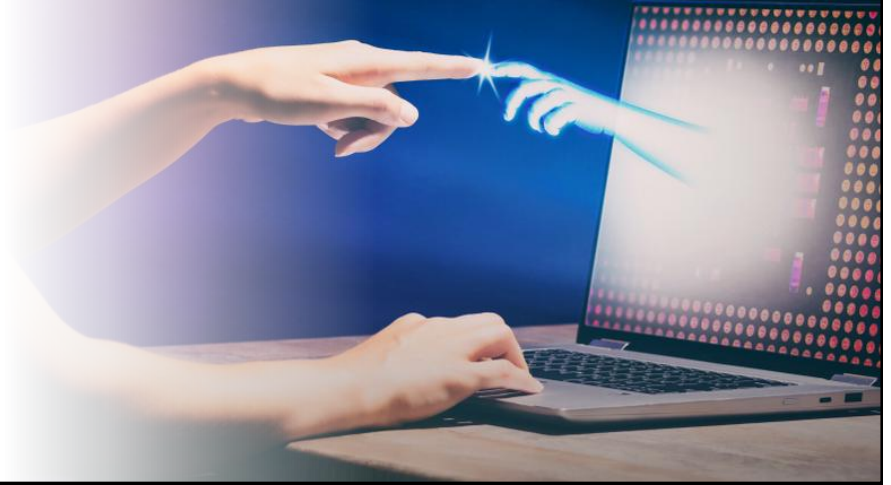
## Framework Analysis (Ritchie et al 1994)

1. Familiarise
2. Develop theoretical framework
3. Indexing (coding)
4. Charting
5. Synthesizing



25

## How are themes generated?



26

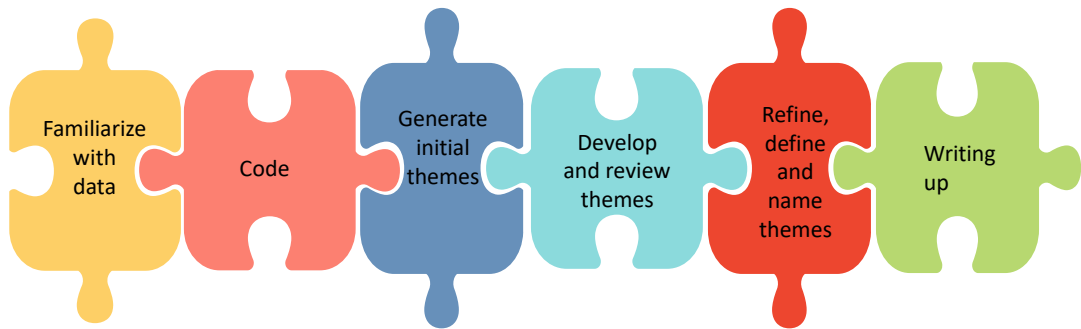
## Human-generated themes

"Patterns are generated **by the researcher** through a rigorous process of data familiarisation, data coding, and theme development and revision. The method can be and is applied in lots of different ways, to lots of different datasets, to address lots of different research questions, and within a range of theoretical frameworks!"

**<https://www.thematicanalysis.net/understanding-ta/>**

27

# Contribution of CAQDAS tools



28

# AI-generated themes <https://www.ailyze.com/faqs>

How does AILYZE identify themes in the data?

Answer: AILYZE uses an inductive approach to theme identification by analyzing the entire dataset. It focusses not only on commonly discussed themes but also unique and minority viewpoints that are noteworthy.

How does AILYZE handle contradictory themes?

Answer: Rather than dismissing or simplifying these conflicts, AILYZE highlights them as crucial for understanding the complexity of the dataset. This method allows AILYZE to present a nuanced view that includes both dominant and exceptional opinions.

Can AILYZE generate codes for the themes?

Answer: No, because it does not using coding. AILYZE uses an inductive approach to theme identification by analyzing the entire dataset. This avoids the limitations of traditional coding methods that focus on single responses and may overlook broader themes or complex discussions. This also ensures consistent theme detection across various segments of data.

29

# Auto-extraction in Gen-AI Aps

## AILYZE

- AI generated themes
- User-generated themes
- Frequency analysis based on AI or user themes
- Shared and unique viewpoints

## Reveal


- Identifies themes in form of topics from interviews
- Similarities and Differences
- Observations Map = thematic visualisation of themes and their presence

## Tailwind (by MAXQDA)

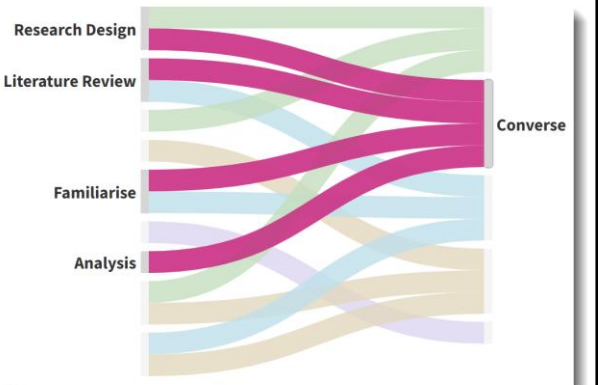
- Identifies themes in form of topics from text files
- Adjustable topic names/descriptions
- Summary Table comparisons


Auto-extraction can be focused by prompting

30



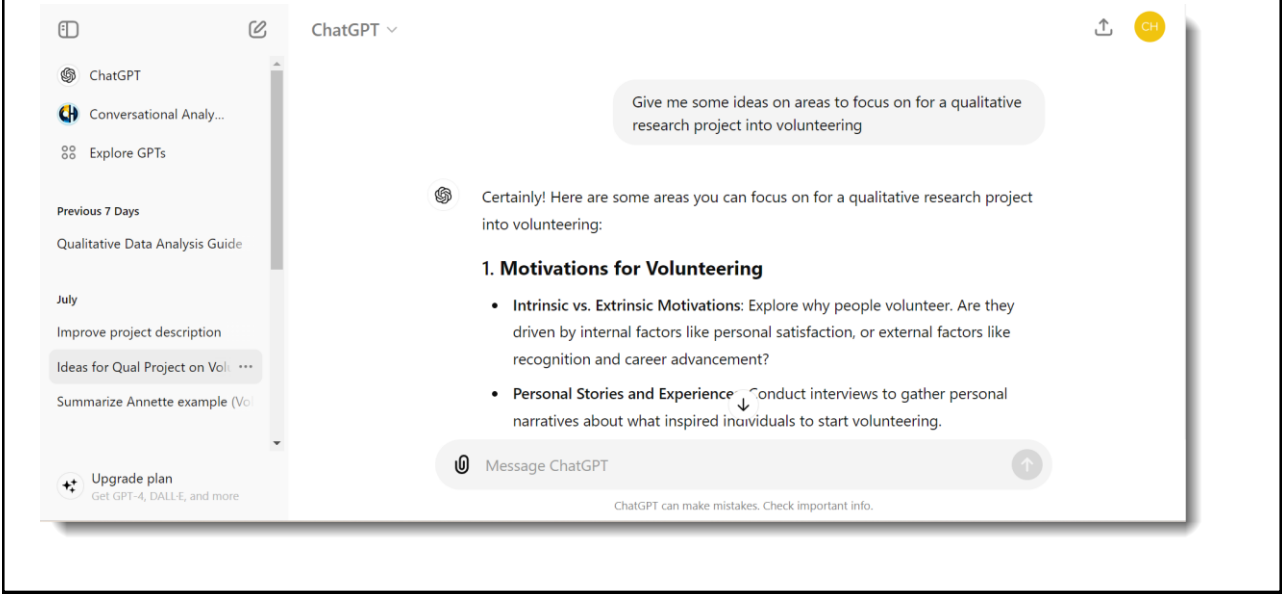
# Conversing: Chatbots for QDA





34

# Intuitive interface ... believable content



35

# Prompting is key

“if you are willing to take the output text extruded by these machines as the answer to what you're looking for, and you can shape the question so that you can make sense of it that way, then they can do it. In other words, **it's all on the people using it and not on the machine**” **Prof. Emily M. Bender** from *Mystery AI hype theatre 3000*

BECAUSE...

⚠️ AI language models are not trained to give **the answer**. They are trained to give a **likely possible answer** that you as a researcher can then check using the evidence provided.

We'll look at examples later

[CoLoop Documentation](#)

36

## Overview of conversing

### ATLAS.ti

- Chat with one document
- Chat with multiple documents

### MAXQDA

- Chat with one document
- Chat with coded-segments

### ChatGPT

- Ask questions of what you upload (note there are limits)
- Use custom GPTs – more on this later

Currently this is max 128.000 characters (approx. 3 or 4 hour long interviews)



ChatGPT

how many interview transcripts can I upload at one time for you to analyse?



You can upload as many interview transcripts as you'd like, but it's generally more practical to handle them in smaller batches for effective analysis. I can process a large volume of text, but breaking it down helps in providing detailed insights, ensuring nothing is missed.

If you're uploading text files, the content should ideally be divided into manageable sections or files, so I can focus on analyzing specific aspects or themes within each batch. If you have a large number of transcripts, you might consider grouping them by topic, interviewee, or any other relevant category.

Feel free to start with a few transcripts, and we can go from there!

In addition is the context window limit – the maximum number of tokens the LLM can process in one go

<https://zapier.com/blog/context-window/#what>

37

## Prompts must be...

- Clear, detailed, specific
- Related to the analysis objectives
- LLMs not very good (yet) at generating quantitative information (this is improving...)

38

# Examples of good/bad prompts

(from the CoLoop documentation)

| Tip                           | Good prompt   | Bad prompt   |
|-------------------------------|---|--|
| Provide context               | Summarise what voters said about their first time voting experience | Summarise the main points                            |
| Communicate in the imperative | Format your last message into bullet points                         | Can you format your last message into bullet points? |
| Use specific language         | What did participants say they liked about concepts A, B and C?     | What did participants say they liked?                |
| Break questions up            | What are some of the pros of concept A?                             | What are the pros and cons of concept A?             |

39



## Resources on the Qeludra Community

<https://qeludra.com/community>



RESOURCES

### Prompts for Qualitative Research

Prompts that will help you explore the use of AI in qualitative research.

40

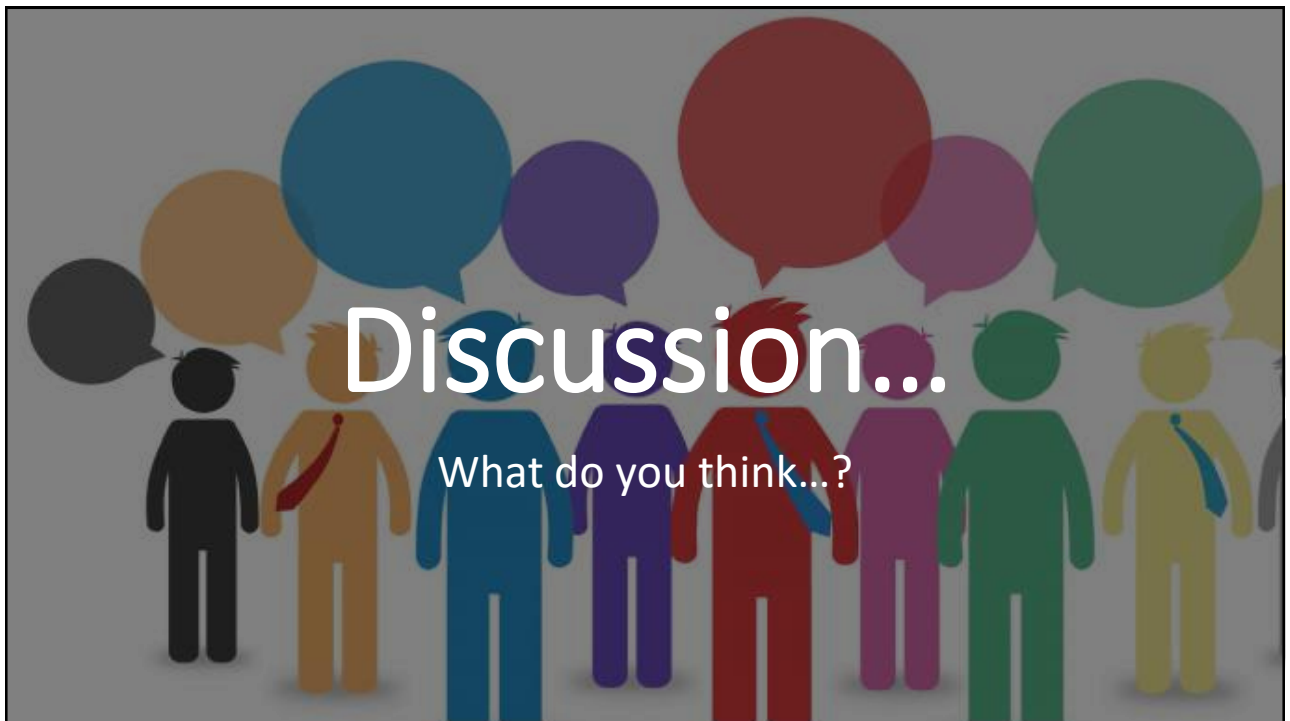


Christina Silver, PhD  
christina@qdas.co.uk

## Query-Based Analysis (Morgan 2023)

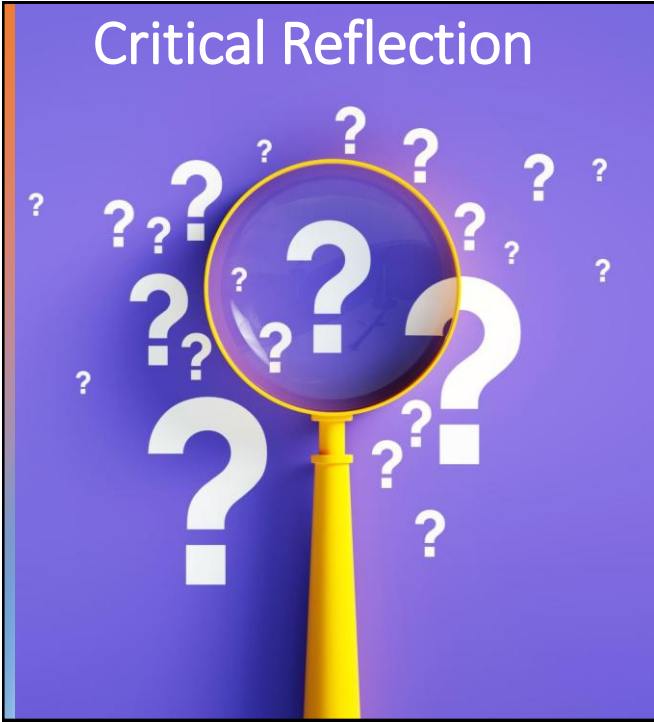
1. Asking Broad, Undirected Queries
  - To locate a set of basic concepts that serve as a foundation for further querying (predicated by context-setting)
2. Following Up with More Specific Queries
  - To generate details about concepts identified in Step 1 (i.e. like sub-categories if you were coding)
3. Examining the Supporting Data
  - To substantiate the concepts identified in Steps 1 and 2

43



44


# Critical Reflection




- Why
- When
- How
- What
- Does

45

# What is...



Gained



Lost

46