Conversational Analysis with AI - CA to the Power of AI: Rethinking Coding in Qualitative Analysis

Susanne Friese, Max-Planck-Society Göttingen, Germany

Abstract

The rapid emergence of generative AI tools challenges traditional assumptions about qualitative data analysis, particularly the central role of coding. This article introduces Conversational Analysis to the Power of AI (CA^{AI}), a novel methodological framework that replaces coding with structured, dialogic interaction between researchers and large language models. CAAI reimagines analysis as a process of iterative questioning, synthesis, and reflexive interpretation rather than segmentation and categorization. Grounded in a hermeneutic epistemology and emphasizing methodological rigor, CA^{AI} integrates inductive, deductive, and abductive reasoning strategies. It allows researchers to adapt procedures from established methods like Grounded Theory while embracing a distributed and co-constructive model of knowledge creation. The article outlines a five-step process for CA^{AI}, discusses reliability and validity in this new paradigm, and positions the approach within broader shifts toward post-coding qualitative inquiry. CA^{AI} offers a compelling alternative for researchers seeking to deepen interpretation, democratize analytic access, and expand the epistemic horizons of qualitative research in the age of AI.

Introduction

The rapid emergence of generative artificial intelligence (genAI) tools is profoundly reshaping how researchers engage with unstructured data, particularly within qualitative analysis. While these advancements offer exciting possibilities, they also bring challenges. Traditional qualitative coding methods, though rigorous, are often time-intensive. Conversely, while general-purpose chatbots and large language models (LLMs) offer speed, they frequently fall short in providing the necessary rigor, transparency, and traceability crucial for sound qualitative inquiry. Many current applications claiming to automate qualitative analysis often perform tasks closer to classification proxies rather than engaging in the deep, interpretive work characteristic of the field.

This article questions the continued necessity of the traditional coding paradigm in this new technological era. It argues that coding, once an essential cornerstone of qualitative analysis, may now be effectively replaced by AI-supported data retrieval combined with dialogue-based interpretation. To this end, I introduce and detail a novel method: Conversational Analysis to the power of AI (CA^{AI}), proposing a shift towards interactive engagement with qualitative data. Furthermore, the article delves into the fundamental epistemological questions surrounding how knowledge is constructed when AI becomes part of the analytic process and proposes concrete strategies for establishing reliability and validity within this new dialogic framework.

Qualitative Coding vs. LLM-Driven Classifications

As large language models (LLMs) become more common in qualitative research, it is increasingly important to distinguish between two fundamentally different approaches to textual analysis with AI: detailed qualitative coding performed by a human, and what might more accurately be called classification proxies. The latter refers to labels assigned (often by AI or automated tools) that resemble qualitative codes but are, in fact, surface-level classifications lacking interpretive depth. While many recent studies claim to automate or accelerate qualitative analysis using generative AI, a closer examination reveals that they often bypass the very practices that define interpretive, in-depth qualitative work—a tendency that is closely tied to the institutional and disciplinary backgrounds of their authors.

Papers authored by researchers from computer science, data science, or human-computer interaction (HCI) departments (e.g., Zhang et al., 2024; Gao et al., 2023; Xiao et al., 2023; Zackary, 2024) typically adopt an engineering mindset—treating qualitative data as text to be segmented, labelled, or clustered using LLMs. These projects frequently rely on custom tools, API pipelines, and pre-processed datasets, often reducing analysis to a set of classification tasks, which requires chunking the text. API (Application Programming Interface) are a set of rules that allows different software systems to communicate with each other. Chunking, however, destroys the continuity that is essential for making interpretive leaps in qualitative analysis. In addition, the number of codes applied is mostly small (usually under 10), and the data sets—while sometimes described as "large"—consist mostly of short texts like survey responses or tweets.

Traditional qualitative coding is not simply the act of assigning labels to text. It is an iterative, interpretive, and contextually grounded process that unfolds over time. Researchers typically engage with full transcripts—often ranging from 10 to 35 interviews—developing between 80 and 250 distinct codes, frequently arranged into hierarchical categories and subcodes. What distinguishes this process is not just the volume or granularity of codes, but the emergence of insight through repeated immersion, reflection, memoing, and reconfiguration.

Qualitative coding is also deeply dialogic and theory-informed. Analysts shift between inductive openness and deductive structure, exploring contradictions, emotional tone, and the subtle meaning behind participants' words. The goal is rarely just thematic description, but interpretive depth, often culminating in the development of conceptual categories, metaphors, or grounded theories.

In contrast, many studies claiming to perform "qualitative analysis" with LLMs (e.g., GPT-3.5, GPT-4, Claude 2) are actually conducting shallow classification over small or pre-segmented datasets. For example, Zackary (2024) had ChatGPT code 232 pre-defined text passages using just 9 codes in 3 categories—a setup that bears closer resemblance to quantitative content analysis than to qualitative inquiry. Similarly, Xiao et al. (2023) applied 9 predefined categories to 668 survey questions. While these tasks are framed as "qualitative coding," they lack the scope, complexity, and context integration typical of real coding practice.

Even when studies use inductive approaches, such as in Deiner et al. (2024) or Perkins and Roe (2024), LLMs tend to produce a small number of high-level themes—often five to ten—with little depth or subcoding. What's more, these outputs are often non-reproducible, context-blind, or overly reliant on surface-level phrasing. This tendency is amplified by the inability of many systems to maintain memory across large document sets.

What much of this emerging literature reflects is not the automation of qualitative coding, but a redefinition of the task to fit the model's limitations. Rather than expanding the epistemic reach of qualitative methods, LLMs are often used to simplify or flatten them into labelling exercises, where meaning is predefined or extracted from context-free units. This approach may generate speed, but it sacrifices reflexivity, nuance, and methodological rigor.

By contrast, studies authored by qualitative researchers from fields like sociology, education, health, or communication (e.g., Hayes, 2025; Goyanes et al., 2024; Lixandru, 2024; Hitch, 2024; Perkins & Roe, 2024) tend to emphasize interpretation, reflexivity, and transparency.

These authors often work directly with interview transcripts or focus groups, engage with models through dialogue, and reflect critically on the limits and affordances of generative AI. Thus, rather than framing LLMs as tools that replace qualitative analysis, it is more productive to see them as co-analysts, agents that support reflection, pattern recognition, or theory generation when guided appropriately. As Hayes (2025) and Ibrahim and Voyer (2024) suggest, meaningful use of generative AI requires dialogic interaction, theoretical framing, and critical engagement. Models like GPT-4 and Claude 2 may be powerful collaborators, but their usefulness hinges on human researchers remaining firmly in control.

In sum, the promise of LLMs in qualitative research lies not in mimicking coding software, but in supporting new methods of inquiry that preserve the interpretive richness of the field. Recognizing the difference between classification proxies and 'true' qualitative coding is essential if we are to ensure methodological integrity in this rapidly evolving space.

Hybrid Solutions: Traditional Coding Tools Enhanced with Generative AI

As generative AI (GenAI) continues to evolve, its integration into established computer-assisted qualitative data analysis software (CAQDAS) marks a significant development in the qualitative research landscape. Rather than disrupting existing paradigms, tools like MAXQDA, ATLAS.ti and NVivo have introduced GenAI features to enhance rather than replace traditional coding-based workflows. These hybrid solutions offer methodological continuity while increasing analytic efficiency, particularly in early-stage tasks such as data familiarization, code frame development, and results synthesis.

MAXQDA's AI Assist provides a representative example of this trend. Unlike general-purpose chatbots, AI Assist is embedded in a rigorously structured environment aligned with formal coding methodologies, such as Schreier's Qualitative Content Analysis (QCA) framework (Schreier, 2012). It supports both concept-driven and data-driven coding strategies, offering targeted assistance that aligns with the logic of the underlying method. When using a concept-driven approach, researchers can define a code, add a description, and have AI Assist apply it segment by segment—document by document. Each application includes an explanation of why the code was assigned, enabling transparency and offering opportunities for critical reflection. This process has also proven useful when treating the AI as a second coder, especially during pilot phases to test and refine category definitions.

For data-driven coding, MAXQDA's AI Assist can suggest paraphrases for individual segments, generate initial concept labels, or propose subcategories under existing codes. While this support is more limited in scope and speed—especially since responses are generated one at a time and do not retain memory across interactions—it still serves as a helpful brainstorming aid. Researchers remain firmly in control of analytic decisions, treating AI output as inspiration rather than automation.

ATLAS.ti takes a markedly different approach by offering full-scale, automated AI-powered coding. However, its implementation raises important concerns about methodological coherence. Unlike MAXQDA's step-by-step, researcher-guided interaction, ATLAS.ti allows users to select all documents for analysis at once—but instead of analyzing documents holistically, the AI processes and labels each paragraph in isolation. It neither tracks previously used codes nor aggregates meaning across a document or dataset. This leads to severe code proliferation: upwards of four hundred codes for two documents, over six hundred for four, and more than 1,200 for just ten, creating an overwhelming volume of redundant or overly granular codes. The time spent cleaning and consolidating these outputs often exceeds the time supposedly saved. While the AI summarization function is relatively well implemented, the coding logic appears driven more by technical affordances than by qualitative principles. The result is a system that exemplifies what can go wrong when AI is integrated without a deep understanding of interpretive methods or the analytical discipline coding demands.

Shortly after the first gen-AI features appeared in CAQDAS platforms, standalone apps built entirely around LLM logic began to emerge—and that's when I first articulated my premise that coding might become obsolete in the age of conversational AI. In my blog post "Rethinking Qualitative Data Analysis: Do We Truly Want a Faster Horse?" (Friese, 2023a) and in the talk "Life Without Coding" (Friese, 2023b), I argued that real-time, dialogic engagement with unstructured data offers a far more natural and flexible alternative to traditional code-and-retrieve workflows. I demonstrated this approach with an example in "Ethical and Responsible Use of AI for Qualitative Analysis" (Friese, 2024) and first started to describe it in Friese (2025). What once seemed a radical departure is now quietly influencing legacy platforms: features like MAXQDA's Tailwind—which let users explore data without applying formal codes—signal a growing acceptance that rigorous qualitative insight needn't rely solely on discrete, manually applied codes.

Indeed, with the rise of retrieval-augmented generation, semantic search, and conversational AI, we are witnessing the emergence of new paradigms for working with unstructured data. Retrieval-augmented generation (RAG) combines document retrieval with language generation, allowing AI systems to ground their responses in relevant source material rather than relying solely on pre-trained knowledge. Coding, long treated as a necessary intermediary for enabling analytic access, may no longer be essential. Instead, we are beginning to see how real-time interaction with data can enable more flexible, layered, and responsive inquiry. Rather than assigning static codes, researchers can engage in dialogue with their material, iteratively probing meaning and retrieving nuanced responses in context.

This shift opens profound epistemological questions—not only about *how* we analyze qualitative data without coding, but also about *what it means to know* in this new paradigm. If coding is no longer the default analytic mode, how do we conceptualize the process of knowledge creation in qualitative research?

What It Means to Know: Epistemological Shifts in AI-Assisted Qualitative Analysis

Traditionally, qualitative inquiry has been grounded in the interpretive paradigm, where knowledge is constructed through the immersive engagement of researchers with empirical material, situated within social contexts and informed by theoretical perspectives. Meaning does not reside in the data itself but is generated by a recursive relationship between the analyst, the context, and the text. In this view, knowing is deeply shaped by the knower's positionality, experience, and reflexive engagement.

As Krähnke, Pehl, and Dresing (2025) argue, a hermeneutic epistemology is particularly well-suited to conceptualizing how AI might be meaningfully integrated into qualitative analysis. In this view, knowledge is constructed dialogically—through the interaction between researcher, text, and interpretive horizons. When AI is added to the equation, it does not replace the researcher but becomes part of a triadic interpretive space, one in which the LLM's outputs act as provocations that the researcher responds to, questions, and situates. The act of understanding becomes a process of navigating between perspectives: that of the participant, the AI model, and the researcher's own evolving framework. The hermeneutic circle is not broken but widened.

The AI model becomes a new epistemic actor that simulates understanding through probabilistic modeling. Even though these models lack *Seinsverbundenheit*—the ontological embeddedness in lived, socio-historical worlds that informs human interpretation (Mannheim, 1936; Gadamer, 1960). Thus, they do not *experience*, *feel*, or *reflect* in the ways human researchers do. Yet, paradoxically, their output can still offer viable insights, precisely because they are trained on vast and socially-situated corpora.

This raises a fundamental epistemological question: if AI-generated interpretations are not grounded in human consciousness or lived experience, can they nevertheless support meaningful qualitative analysis? The answer, as Krähnke et al. (2025) suggest, depends less on what AI is and more on how it is used. Large language models can serve as abductive catalysts—generating unexpected insights, challenging established assumptions, and inviting new interpretive directions. Their outputs do not arise from understanding in a human sense, but from vast intertextual exposure. The digital traces that fuel their training are themselves reflections of the lifeworlds we seek to study—and many more that lie beyond our personal reach. In this sense, the "knowledge" produced by AI is not personal but collective and distributed—an aggregated echo of meaning-making across diverse social, cultural, and discursive contexts.

Engaging with LLMs, then, allows researchers to access perspectives that may otherwise remain inaccessible. These models can introduce interpretive angles, nuances, or contrasts that extend the scope of human inquiry. When used dialogically and iteratively, they become tools for epistemic expansion, offering a form of interpretive triangulation: not between data sources, but between human insight and machine-generated associations. In this configuration, researchers do not cede authority to the model, but remain in control—accepting, rejecting, or refining AI outputs in light of their theoretical, contextual, and ethical frameworks.

These developments compel us to revisit what counts as qualitative knowledge. Must it always stem from human immersion alone, or can it also arise from orchestrated interactions with systems that mirror, provoke, and extend human sense-making? The epistemological terrain is shifting: from individual hermeneutics to distributed intelligence; from solitary interpretation to dialogic co-analysis. In this space, knowing is less about ownership and more about orchestrating meaningful epistemic encounters—between researcher, data, and AI model.

As researchers move toward more dialogic, reflexive forms of AI-assisted inquiry, they are not abandoning qualitative principles but reimagining them. The challenge is not to preserve tradition for its own sake, but to ensure that the emerging practices—however novel—remain grounded in the ethical, contextual, and interpretive commitments that define the field.

Emerging Practices of Post-Coding Analysis

Above, I already challenged the assumption that coding is intrinsic to qualitative analysis and argued that, in light of LLM capabilities, a new mode of inquiry is not only possible but increasingly necessary. How do analysts ensure rigor, reflexivity, and conceptual depth in an environment where categorization is no longer a prerequisite? These are the questions explored in the next section, which shows that coding-free analysis is not merely speculative but already emerging in practice.

A number of researchers are beginning to chart pathways toward post-coding analysis—tentatively at first, often using commercially available tools like ChatGPT. Yet many of these efforts still reflect a partial adherence to the logic of classification that has long structured qualitative research. Hayes (2025) exemplifies a hybrid approach that begins with traditional coding logic but moves decisively toward interactive engagement with AI. His process starts by asking GPT-4 and Claude 3.5 to generate a thematic coding framework based on full interview transcripts and to classify documents according to these themes. For instance, Claude outputs a codebook and a table indicating whether specific codes appear in each document, along with a few illustrative quotes. While this reflects a classification mindset, Hayes does not stop there. The second phase of his analysis shifts toward dialogic exploration. Using the thematic scaffolding as a starting point, Hayes engages the LLMs in more dynamic tasks: he prompts them to compare perspectives across cases, simulate policy debates, or reflect on contradictions in the data. In doing so, he treats the LLM less as a coding engine and more as a conversational partner—albeit one whose outputs are treated with caution and constantly evaluated against the researcher's interpretive framework.

Perkins and Roe (2024) move a step further away from coding. Working with institutional policy documents, they use ChatGPT to suggest high-level themes and then request quotations to support or refine those themes. Unlike Hayes, they do not construct or apply a formal codebook, nor do they attempt comprehensive classification. Instead, the LLM is used for pattern recognition and plausibility checks, with the human researcher validating outputs. Their approach decouples theme development from data segmentation, treating the AI as a thematic assistant, while reasserting human control over interpretation and trustworthiness. The analysis remains descriptive, but their method clearly signals a departure from the rigid mechanics of coding.

Morgan's (2025) Query-Based Analysis (QBA) pushes this line of thought further. He presents a structured, reflexive alternative to traditional coding, grounded in a three-step process that blends inductive logic with iterative dialogue. The approach begins with broad, undirected queries posed to an LLM in chat interface. These initial prompts—formulated with careful contextual framing—are used to elicit a range of high-level themes from the data, serving as entry points for deeper inquiry. Rather than seeking a single definitive answer, Morgan emphasizes the practice of prompt engineering, comparing responses to differently phrased questions and selecting the most meaningful set of candidate themes based on the researcher's familiarity with the data and interpretive goals.

In Step Two, the researcher follows up with more specific queries, targeting each of the previously identified themes to elaborate potential subthemes. These prompts allow for the surfacing of overlapping or interrelated concepts, which are then refined, consolidated, and evaluated by the researcher. This process highlights Morgan's emphasis on reflexivity and analytic judgment—recognizing that subthemes are not fixed categories but evolving conceptual constructs that gain clarity through iterative engagement.

The final step, Step Three, focuses on substantiating themes and subthemes with supporting quotations. While this resembles traditional thematic presentation in qualitative reporting, the path to these findings has been dialogic rather than classificatory. The LLM functions not as a coding engine, but as a responsive partner for exploration—returning to specific passages when queried and allowing researchers to combine theme-based and relational searches. QBA therefore offers a compelling post-coding framework: one that retains structure and transparency yet dispenses with segmentation and labelling as a prerequisite for analysis. It positions the LLM as an iterative thinking tool, with the human researcher maintaining epistemic authority throughout. While the analysis leans descriptive, QBA provides a lightweight, adaptable alternative to manual coding, especially valuable for researchers seeking analytic traction without the burden of segmentation or labelling. It signals a growing recognition that *interacting with* data can be just as productive as *categorizing* it.

Building on this shift, Nguyen-Trung and Nguyen's (2025) *Narrative-Integrated Thematic Analysis* (NITA) offers a more developed and structured analytic framework that moves beyond coding without fully leaving behind thematic traditions. Rather than segmenting data into coded units, the analysis process begins with the identification of high-level themes across interviews, followed by the construction of individual narrative profiles for each respondent. These profiles are then compared and synthesized into cross-case narratives, allowing the AI to surface commonalities, contrasts, and recurring motifs.

Throughout the process, the researcher remains in control—curating AI suggestions, refining theme boundaries, and evaluating the coherence of emergent storylines. The AI is not treated as an authority, but as a reflective collaborator that can offer plausible interpretations, test comparisons, and challenge assumptions when prompted effectively. The result is a form of synthesis that retains the structural logic of thematic analysis while being freed from manual coding and categorical rigidity.

NITA reflects a transitional model: one foot in the familiar world of theme development and the other stepping toward a new paradigm of AI-assisted narrative construction. It acknowledges that interpretive insight can emerge not from labelling discrete data fragments, but from engaging with participants' accounts as whole, evolving narratives—especially when supported by a language model configured to work dialogically.

Together, these approaches suggest that qualitative research is entering a moment of methodological experimentation. While none of the above authors propose fully codified alternatives to the traditional paradigm, their work points to the same core insight: that dialogue—not classification—may become the foundation of future analysis.

Building on these emerging practices, Conversational Analysis with AI (CA^{AI} – read: *CA to the power of AI*) takes the next step: not adapting qualitative analysis to fit the logic of LLMs but reimagining the analytic process itself around human–AI dialogue.

CA to the Power of AI: A Method for Dialogic Analysis in the Age of LLMs

CA to the power of AI (CA^{AI}) responds to the growing recognition that coding may no longer be necessary as the backbone of qualitative analysis. It offers a clear proposition: that a rich and rigorous analysis can emerge through sustained dialogue between human researcher and AI model—without ever applying a single code. Rather than framing LLMs as tools for accelerating coding, this method sees them as interactive sense-making partners—ones that retrieve, synthesize, and contrast segments of data based on researcher-led questioning. It is grounded in the principle that understanding does not emerge from labelling, but from the layered exploration of meaning across context, narrative, and abstraction. While the examples provided below use interview transcripts, the dialogic principles of CA^{AI} are equally applicable to other forms of unstructured qualitative data, such as focus group discussions, observational field notes, or documents.

The method unfolds in four iterative steps, with an optional fifth for those seeking to move into theory-building. Each step positions the researcher as an active analyst who guides the AI toward insight—not by delegating analysis, but by cultivating it through a structured process of inquiry and reflection. To illustrate the method in practice, each step is accompanied by an example drawn from my PhD thesis, which examined the relationship between impulse buying and self-identity across three shopper groups: addicted buyers, compensatory buyers, and utilitarian buyers. The original study comprised a total of 57 interviews, for this article only 9 of the 57 interviews are used—3 from each shopper group—because of space reasons. The analysis was conducted using QInsights, an AI-powered tool developed specifically for dialogic qualitative analysis (QInsights BV, 2025).

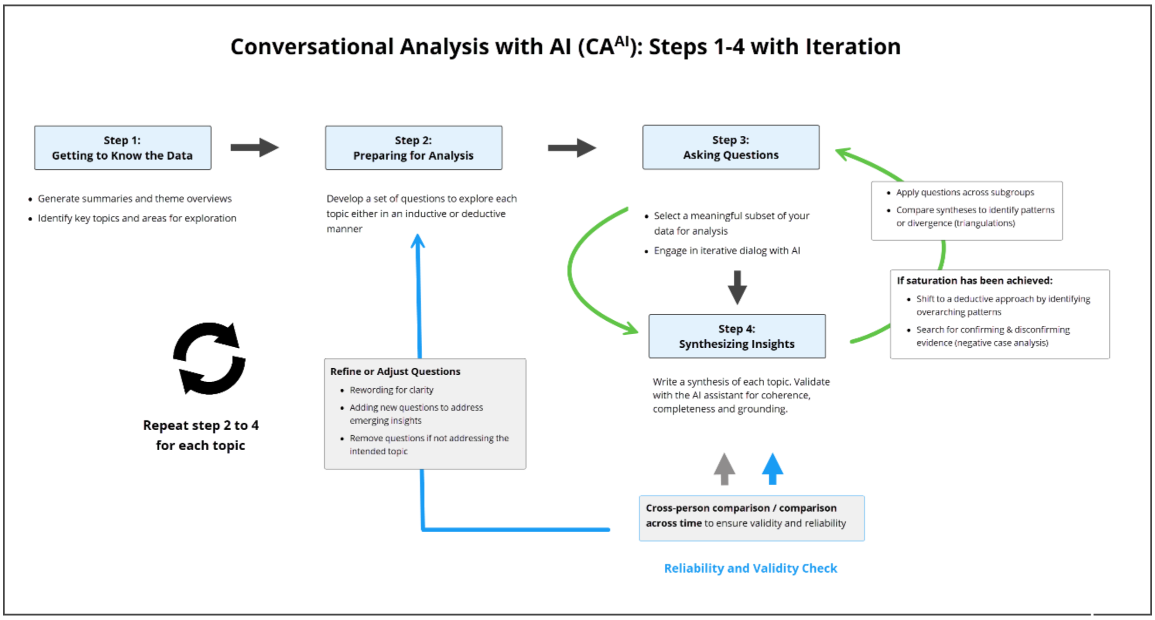

Figure 1: Step 1 to 4 of Conversational Analysis with AI

Step 1: Getting to Know the Data

Researchers begin by generating summaries and extracting preliminary themes with the help of an AI assistant. This serves as an initial orientation—identifying key areas for further exploration. In content analysis, these themes may align closely with research questions; in a grounded theory-style approach, a single emergent theme might be used to initiate inquiry and evolve iteratively.

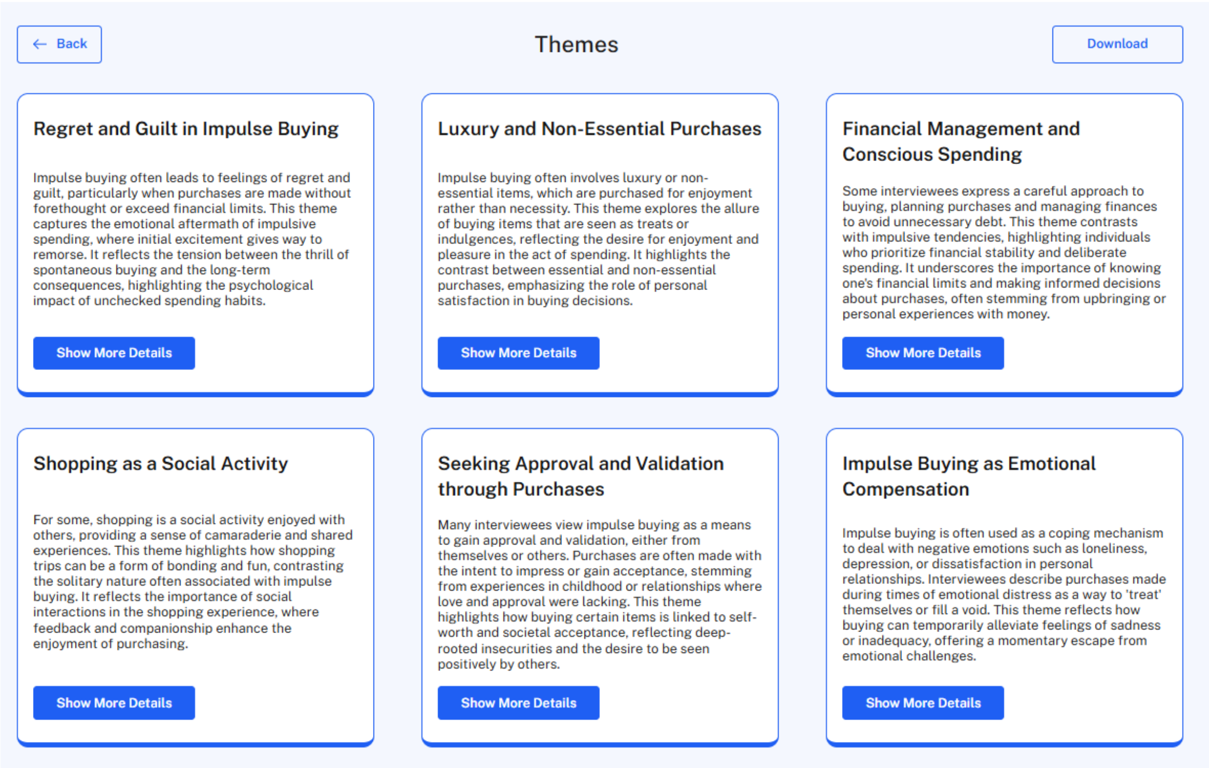

Figure 2: Suggested themes (Screenshot: QInsights Themes Analysis)

An example prompt to extract themes could look like this:

Analyse the provided qualitative data. Identify and list between 2 and 10 recurring themes. Each theme should be distinct and not overlap with others. For each identified theme, provide the following information:

1. A concise, descriptive title for the theme.

2. A detailed description of the theme (around 100 words).

Step 2: Preparing for Analysis

Researchers select one topic to start with and develop a set of exploratory questions to guide the interaction with the LLM. These questions may be inductive, open-ended prompts that allow patterns to emerge organically from the data, or deductive, shaped by existing theoretical frameworks or prior findings. In both cases, the aim is to surface relevant data segments and support meaningful interpretation aligned with the researcher's analytic intent.

While it can be helpful to brainstorm the questions with the AI assistant, the final selection and wording of questions is firmly in the hands of the researcher. This question set becomes the analytical scaffolding for the topic and also serves as a means to make the analysis transparent and replicable. Thus, CA^{AI} analysis proceeds topic by topic. This contrasts with conventional coding frame development, where researchers typically seek diversity early on, coding documents sequentially across all topics that occur.

Continuing with the example, the topic selected for analysis is one of the above identified themes: *Seeking approval and validation through purchase*. To explore this topic, the following set of guiding questions was developed:

- What explanations do respondents offer for making purchases they later regretted or found excessive?
- How do respondents describe the emotional state or life situation they were in before and after these purchases?
- In what ways do respondents connect their buying decisions to social expectations or a desire for external validation?

I typically conclude the dialog with a set of follow-up prompts that ask the LLM to summarize each case, highlight similarities and differences between respondents, and identify any unique or divergent perspectives. Instead of a case-based summary, you can also request a summary that integrates perspectives across all respondents.

- Write a summary in essay format about (respondent name), focusing on her motivations for impulse purchases, her emotional states before and after buying, and any connections to social expectations or the desire for validation, as discussed in our questions. Include direct quotes with page numbers to support key points and maintain the respondent's authentic voice.
- How do these respondents differ, and what do they have in common?
- Do any respondents offer a unique or divergent perspective?

Step 3: Asking Questions

While it is technically possible to process 20 or more one-hour interview transcripts at once, doing so tends to result in generalized responses. Generative AI models are prone to summarizing rather than analysing when faced with large-scale input, which can smooth over important nuances. For this reason, it is best to begin with a focused subset of the data—typically 4 to 6 interviews, and rarely more than 10 at a time, depending on length and complexity. Starting small allows the researcher to stay close to the data and retain greater control over the analytic process.

In this step, the researcher engages the AI Assistant in a dialogic exchange using the questions developed in Step 2. AI responses are not treated as definitive answers but as openings for deeper exploration—prompting follow-up questions, requests for clarification, comparisons across cases, or challenges to initial impressions. Through this interactive back-and-forth, interpretation begins to take shape—not as something delivered by the AI, but as a process the researcher actively guides and participates in. Figure 3 illustrates such a dialogue. Its purpose is to highlight the structure of the interaction rather than the specific content.

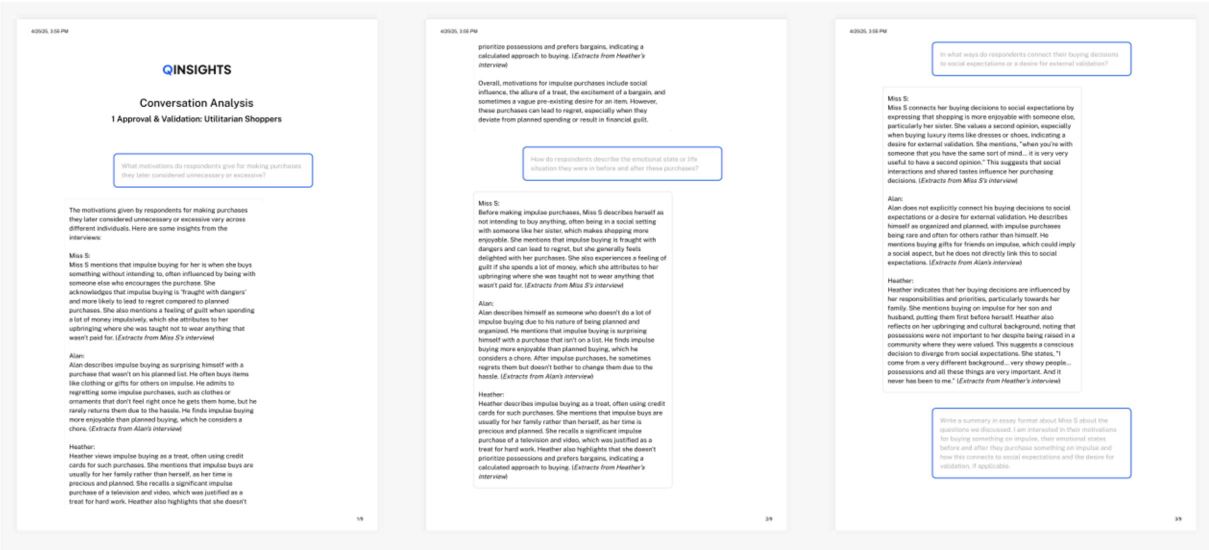

Figure 3: Chat protocol serving for later synthesis, also adding transparency to the analysis process

Typically, researchers engage in **inductive** or **deductive** reasoning during this step, depending on how the initial questions were framed. However, when unexpected findings arise—observations that challenge assumptions or fall outside existing conceptual frameworks—**abductive reasoning** becomes particularly valuable. First described by Charles Sanders Peirce, abduction refers to the process of developing the most plausible explanations for puzzling or surprising observations. Researchers may notice contradictions, outliers, or unexpected statements—either intentionally sought or discovered serendipitously—which then lead to the formulation of tentative hypotheses. Here, the AI Assistant can become a creative thinking partner in the abductive process (see also: Variations to the Analytic Dialog). Researchers can:

- Ask exploratory questions to probe anomalies (e.g., Why might respondents with similar experiences express contrasting emotions?)
- Brainstorm plausible explanations, drawing from cultural, psychological, organizational, or contextual perspectives
- Iterate between data and hypotheses to refine, challenge, or expand emerging ideas
- Compare across subgroups to explore whether patterns hold across different contexts

This dialogic exchange—whether guided by inductive, deductive, or abductive reasoning—gradually shapes interpretation. The researcher remains in control, using the AI to broaden perspectives, retrieve relevant data segments, and test emerging insights. Once a topic has been explored in sufficient depth, the process moves on to step 4: Synthesizing Insights.

Step 4: Synthesizing Insights

When you ask an AI assistant questions about your data, it can generate pages of output in minutes. In the case of the *Approval & Validation through Shopping* dataset, the dialogic exploration across the three shopper types produced over 30 pages of text. It's like traveling by high-speed train: you cover a lot of ground quickly, but insight doesn't emerge at that speed. Your understanding—your capacity to reflect, synthesize, and integrate—needs time to catch up. This is where **Step 4** comes in.

After completing the dialogic exploration in Step 3, the researcher slows down to engage deeply with the conversation(s). In QInsights, this can be done directly in the analysis archive or by exporting the exchange. I chose to export the material and read through it in a Word document—highlighting passages, writing comments in the margins, and noting emerging patterns. I returned to the AI conversations when I needed clarification or wanted to probe further. At times, I revisited the original source documents to re-immerse myself in the data and contextualize insights.

From here, there are two valid paths:
You may choose to write the synthesis yourself, grounding your interpretation in theory, experience, and contextual understanding. Or, you may prefer to co-develop the synthesis with the AI Assistant, by uploading the conversation as a new document and shaping a draft collaboratively.

In this dialogic mode, the researcher guides the AI—not only in *what* to say, but in *how* to say it: what aspects to emphasize, which lens to apply, and how to structure the narrative. The AI responds by recombining and reformulating the content based on this guidance, surfacing language or framings that the researcher can accept, reject, or revise. You might also challenge the model to identify weak spots, test counter-interpretations, or offer alternative perspectives.

In this way, the synthesis is neither a product of pure human reasoning nor a machine-generated artifact, but the outcome of structured, intentional interaction. The LLM acts as an epistemic actor, not because it understands, but because it contributes distributed patterns of association—drawn from its intertextual training—that expand the researcher's own interpretive lens.

Insight, then, arises from the interplay between three agents: the data, the model, and the researcher, each shaping and reshaping the other through recursive dialogue. Importantly, this does not mean the researcher cedes interpretive authority. On the contrary, the strength of the synthesis depends on the reflexivity, methodological awareness, and theoretical grounding with which the researcher orchestrates the process. Whether the synthesis is written independently or co-constructed with the model, what matters most is the epistemic choreography: the thoughtful, iterative dance between questioning, interpreting, and refining meaning.

Example Analysis

Below is a potential outcome from the example analysis. The table compares the three shopper groups and highlights key components related to the three guiding questions. This is followed by a summary essay that synthesizes the cross-group insights, exploring the underlying motivations behind impulse purchases, the emotional states before and after buying, and the role of social expectations and external validation in shaping each group's behaviour.

Table 1: Question series 1 - Approval seeking and external validation through impulse buying comparing three shopper types

Aspect	**Utilitarian Shoppers**	**Compensatory Shoppers**	**Addicted Shoppers**
Primary Motivations	Treats, bargains, social encouragement, occasional surprise, gifts for others	Mood repair, desire to appear right, emotional boredom, personal expression, habit	Emotional compensation, validation, control, self-worth, personal history
Emotional State Before Purchase	Neutral or planned mindset; open to suggestion; social context may create a buying mood	Low mood, feeling fed up lonely, off-routine, bored	Depressed, angry, anxious, emotionally neglected, searching for escape or meaning
Emotional State After Purchase	Mild regret or guilt enjoyment, satisfaction	Initial excitement, satisfaction, later regret	High during purchase; guilt, shame, emotional 'hangover' after
Role of Social Expectations / Validation	influenced by sister; influenced by family role; minimally affected	appearance, attention; self-expression	Very strong: buying to impress, to gain approval, to prove worth
Typical Impulse Purchases	Clothing, gifts, electronics (TV), ornaments	Unique clothes, ties, tablecloths, packaged items, ornaments	Home items, high-end clothes, children's clothing, symbolic items of status

Impulse Buying Across Shopper Types: A Comparative Analysis

Impulse buying may seem like a singular behaviour, but when viewed through the lens of different shopper identities—utilitarian, compensatory, and addicted—it reveals distinct emotional drivers, decision patterns, and social influences.

Utilitarian shoppers typically approach shopping with a plan and a purpose. Their impulse purchases are relatively infrequent and often tied to specific contexts: treats for themselves or others, bargains, or social interactions. Miss S, for example, enjoys shopping with her sister and values a second opinion, while Heather frames impulse purchases as rewards for hard work. Though not emotionally driven in the traditional sense, these buyers still experience post-purchase emotions ranging from delight and satisfaction to mild regret or guilt, particularly when purchases deviate from financial norms ingrained by upbringing.

In contrast, compensatory shoppers turn to impulse buying as a form of emotional self-regulation. Their motivations are rooted in lifting low moods, breaking routine, or maintaining appearances. Betina shops to cheer herself up, Steve buys to project an image of success, and Jo sees shopping as a playful escape. Emotional fluctuations are more pronounced in this group: purchases often follow moments of boredom, loneliness, or insecurity, and are accompanied by fleeting excitement or satisfaction followed by occasional regret. Social expectations play a prominent role for Betina and Steve, whereas Jo's purchases are more about self-expression than external validation.

Addicted shoppers, experience the most emotionally intense and cyclical relationship with impulse buying. For them, shopping serves as a coping mechanism for deep-seated emotional needs ranging from unhappy relationships and past neglect to anger or anxiety. These purchases bring an emotional high that temporarily fills a void, but are inevitably followed by guilt, shame, or what Margaret describes as "an alcoholic hangover." Social validation is a dominant theme here: Chris seeks admiration from her husband, Margaret shops to manage how others perceive her, and Donna links quality purchases with her self-worth and social standing. Their impulse buys often carry symbolic meaning: status items, high-quality clothes, or things for their children.

In summary, while utilitarian shoppers occasionally indulge against their better judgment, compensatory shoppers use buying as mood management, and addicted shoppers engage in deeper emotional compensation and validation. The common thread is that impulse buying always says something about the buyer's internal world—whether it's a momentary lapse, a coping mechanism, or a cry for affirmation.

Even though I only analysed 9 documents here—and focused on a single aspect—comparing this experience to the analysis I conducted for my PhD thesis nearly 30 years ago is striking. Back then, I manually coded the full set of 57 interviews using ATLAS.ti version 4. I spent many weeks immersed in coding, followed by additional weeks working through the coded segments to develop my interpretation. For a comparison the figure below shows the categories I have developed that approximate the topics shown in the table. I did not code for the type of items purchased. In total I had around 180 codes.

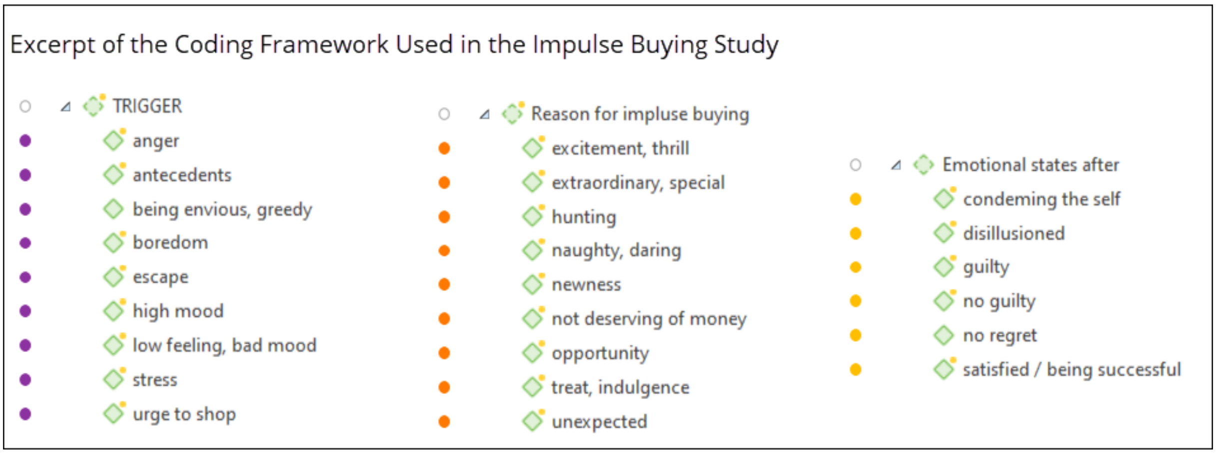

Figure 4: Extract of the coding system from my thesis data

Without doubt, had I worked with an AI assistant in the manner described above, my analysis would have been more thorough, more insightful, and considerably richer and deeper—and all within a significantly shorter timeframe.

Looping Step 3 and 4

When working with subgroups, the same set of questions is applied to each group individually. Each subgroup may reveal distinct nuances or reinforce patterns observed earlier. As the analysis unfolds, new insights accumulate, gradually contributing to a more comprehensive understanding of the topic.

In the example above, I worked with just three interviews per subgroup, so I began preparing the synthesis after exploring all three groups. If you're working with larger subgroups, it can be useful to alternate between Steps 3 and 4: complete the dialogic exploration for one subgroup, write a synthesis, then return to Step 3 for the next subgroup, and so on.

Crucially, the process is iterative, not fixed. While asking questions during Step 3, researchers may discover that something important was left out—or that a new line of inquiry emerges based on the AI's responses. In such cases, the original set of questions from Step 2 can be revised or expanded. These updated questions are then applied in the next loop, allowing the analysis to evolve in response to emerging findings.

When saturation is reached—meaning no new substantial insights emerge—researchers may transition into a more deductive mode in subsequent rounds, testing or refining earlier interpretations against new data. Once all subgroups have been analysed, you can compare the syntheses across groups. To support this final comparison, consider uploading the individual syntheses as a combined document and asking your AI assistant to help identify overarching patterns or differences.

Once a topic has been thoroughly explored across all relevant subgroups, the researcher returns to Step 2 to develop a new set of questions, and the Step 3–4 loop begins again for the next topic.

Ethical and Authorship Considerations

As the boundary between human and machine authorship becomes more fluid, ethical questions emerge around transparency, credit, and responsibility. If an AI model contributes to the generation of analytical text, to what extent should this be acknowledged? Who owns the insight? And how do we ensure that the resulting analysis remains accountable to the standards of qualitative research?

In the CA to the Power of AI approach, the researcher remains the primary author—even when language is co-generated. This is because the AI's contributions are guided, filtered, and ultimately shaped by the researcher's theoretical lens, methodological decisions, and interpretive judgment. The model may suggest formulations, highlight associations, or propose framings, but it is the human who determines what counts as a meaningful insight, and why.

Nevertheless, transparency is essential. In academic contexts, researchers should disclose how AI was used in the analytic process: whether the model was asked to generate summaries, provide quotes, or assist in drafting the synthesis. Rather than undermining rigor, such disclosure strengthens it—making visible the dialogic and iterative nature of meaning-making in this new paradigm.

Ethical use also requires epistemic humility. Researchers must remain aware of the model's limitations: its lack of lived experience, its tendency to reflect dominant discourses, and its potential for hallucination or bias. These risks don't disqualify AI from playing a role in qualitative research—but they do demand that researchers critically interrogate the outputs and make interpretive choices with care and context in mind.

In short, the model may contribute language, patterns, and even inspiration—but the researcher contributes meaning. Authorship lies not in who types the words, but in who decides what they mean and why they matter.

Step 5: Elevating the Analysis

A possible fifth step involves extending the dialogical process between data, researcher, and AI by incorporating explicit theoretical reasoning. At this stage, the objective is not merely to describe patterns but to engage in conceptual abstraction: identifying relationships between themes, integrating findings into existing theoretical frameworks, or contributing to the development of new theoretical insights.

In this phase, the AI assistant remains a valuable partner by helping to surface patterns, suggest theoretical connections, and test emerging interpretations against the data. However, following the quality standards proposed by Strübing et al. (2018), it is the researcher's responsibility to ensure a dynamic, iterative interplay between theory and data—what they describe as "theoretical pervasiveness." Rather than applying theory mechanically or relying solely on AI-driven suggestions, researchers critically engage with both data and theoretical concepts, allowing each to "irritate" and refine the other. Elevating the analysis thus means extending the co-intelligence of researcher and AI into the theoretical dimension—ensuring that findings are not only empirically grounded but also conceptually robust and capable of contributing original insights to the broader scientific conversation.

There are several ways to approach this with the support of an AI assistant. One option is for the researcher to identify and articulate relationships or patterns that seem to emerge across the various topic-level syntheses. These hypotheses can then be tested or explored further in dialogue with the AI by revisiting the data, examining additional examples, or asking targeted follow-up questions. Alternatively, the AI assistant can be prompted to suggest connections across themes, though this typically requires the researcher to provide the relevant syntheses as input, especially if the analysis is being continued in a new session, where prior context may no longer be available in the model's short-term memory.

Another approach is to relate findings to theory. Researchers may ask the AI whether a particular theoretical framework could be relevant to explaining the observed patterns. To avoid unreliable or overly general responses, researchers are advised to either provide a concise summary of the relevant theory or verify the AI's interpretation by asking it to explain the key components. With the theoretical groundwork in place, the AI assistant can then be used to explore how the findings align with the theory, where inconsistencies arise, or how the theory might be extended or adapted in light of the empirical insights.

Step 5 thus supports deeper reflection and meaning-making—whether through theory-driven interpretation, comparative exploration, or conceptual integration. It marks a shift from topic-level analysis to broader insight, allowing researchers to position their findings within broader scholarly or strategic frameworks.

Variations to the Analytical Dialog

To support a reflective and adaptive use of *CA to the Power of AI*, researchers are encouraged to adopt a flexible stance toward their interaction with the language model—not as a static user, but as an active, interpretive agent. Drawing on Thominet et al.'s (2024) conversational roles framework, CA^{AI} invites researchers to consciously shift across their four proposed roles throughout the analytic process. At times, one may act as a *Manager,* ensuring that AI outputs are grounded, accurate, and coherent with the dataset. In another moment, the researcher becomes a *Colleague*, co-theorizing with the model, exploring divergent interpretations, or refining emergent patterns (especially step 3 and 5). The *Teacher* role appears when providing conceptual clarity or steering the model's understanding of terms, cases, or context (e.g. in step 5).

Finally, the Advocate engages the model in imagining alternative perspectives—foregrounding participant voice, marginalized viewpoints, or ethical implications. Here, researchers might draw on strategies proposed by Hayes (2025), such as prompting the AI to take on the voice of a particular stakeholder, generate a persona from an underrepresented group, or simulate a dialogue between competing theoretical frameworks. For example, one might ask: *"How might this finding be interpreted by a first-generation college student?"* or *"Debate this insight from the perspective of feminist versus postcolonial theory."* Such prompts not only diversify interpretive angles but also surface silences and implicit biases, fostering richer, more inclusive interpretations (steps 3 and 5). These roles are not prescriptive but serve as dialogic resources, helping researchers remain reflexively attuned to their shifting positions in a dynamic human–AI analytic encounter.

Adapting Existing Methods within the CAAI Framework

This allows researchers to integrate procedural steps from traditions such as Grounded Theory, Thematic Analysis, or Narrative Inquiry. CA^{AI} serves as a methodological framework, a broader approach that redefines how qualitative analysis can be conducted in partnership with AI. As such, it provides the epistemic scaffolding within which various established methods can be adapted and recontextualized. In this framework, methodological orientation is expressed through the kinds of questions researchers ask and how they structure the analytic conversation with the AI assistant. Inductive strategies invite emergent patterns, deductive ones test pre-defined ideas, and abductive reasoning explores surprising or counter-intuitive insights. The focus shifts from tagging and segmenting to questioning and synthesizing—while retaining rigorous, theory-informed interpretation.

Examples: Grounded Theory, Content Analysis, and Thematic Analysis in CA^{AI}

To illustrate, consider how Grounded Theory procedures might be translated into the CA^{AI} framework. Researchers could begin analysis with the first interview, selecting brief but telling passages for exploration—mirroring *open coding* as for instanced described by Corbin and Strauss (2014). In CA^{AI}, this does not involve assigning codes to segments, but rather "breaking open" the data through exploratory questioning. As provisional concepts emerge, the researcher can guide the AI to retrieve additional examples, examine variation, and organize patterns ultimately supporting conceptual development. At a later stage, what is traditionally referred to as axial coding—*defined as the process of relating categories and concepts through a combination of inductive and deductive thinking*—can also be conducted dialogically, using structured prompts to explore relationships, test assumptions, and refine the analytic structure. No tagging is necessary.

Content analysis can also be adapted within the CAAI framework by rethinking the sequencing of category development. Traditionally, content analysis involves defining categories in advance and fitting data segments into these predetermined slots. In contrast, CAAI turns this process on its head: researchers engage the AI in a dialogic exploration of a topic, using structured questioning to surface the range of aspects, themes, and nuances present in the data.

Rather than coding data to match predefined categories, researchers first build a rich narrative understanding of the topic through dialogue. Only at the end of the process are categories and their dimensions abstracted—grounded directly in the content of the discussions. This approach mirrors the logic illustrated in Table 1, where various aspects surfaced through conversation could later serve as conceptual categories or, in traditional content analysis terms, as codes.

Building on this, thematic analysis within CAAI takes the process one step further. After developing categories and identifying their dimensions, researchers can engage the AI assistant in a new phase of dialogue—searching for overarching themes that connect or transcend individual categories. Through structured, reflective questioning, researchers work with the AI to synthesize broader patterns of meaning, moving beyond description to higher levels of abstraction and interpretation.

Here again, the emphasis is not on mechanical extraction or classification, but on iterative, dialogic sense-making that fosters deeper, more coherent thematic development. Themes are not simply named; they are collaboratively constructed, tested, and refined through analytic conversation. In this way, CAAI offers a pathway for both content and thematic analysis that preserves the integrity and depth of traditional methods, while embracing the flexibility, responsiveness, and epistemic openness of AI-supported dialogic inquiry.

However, while procedural elements can be adapted, the epistemological stance must shift. CA^{AI} assumes a co-constructive view of knowledge that differs from previous methodological assumptions. For instance, while Charmaz's Constructivist Grounded Theory offers an important turn toward subjectivity and reflexivity, it is based on a human-centred model of meaning-making. In contrast, CA^{AI} treats AI as a new kind of epistemic actor—not a conscious knower, but a dialogic partner capable of generating insight through probabilistic modelling trained on socially situated corpora. “Construction” in CA^{AI} is not solely human; it is *distributed*, emerging through interaction between human interpretation, data context, and machine-generated associations. This distinction matters. Adapting methods like Grounded Theory to CA^{AI} is not a matter of transplanting procedures, but of reworking them to reflect a new epistemological logic—one grounded in conversation, co-orchestration, and iterative sense-making across human–AI boundaries.

In sum, CA^{AI} supports methodological pluralism within a unified epistemic frame. It invites researchers to bring their analytic traditions into a new kind of conversation—not only with data, but with a model that extends, challenges, and reframes human understanding. Developing robust question-design practices for different analytic traditions remains an important next step in this evolving field.

Reliability and Validity in CA to the Power of AI

Concerns about reliability and validity—or trustworthiness and rigor, depending on one’s methodological stance—are central to all forms of qualitative inquiry and take on renewed urgency in the context of generative AI. When analysis is fluid, iterative, and co-constructed through dialogue with a language model, familiar markers of rigor such as coding frames or inter-coder agreement do no longer apply. In this shifting landscape, new strategies are required to ensure transparency, consistency, and methodological integrity.

how does this all actually differ from working with a team?

how can you talk about rigor if you start with "what are the themes"

CA^{AI} responds to this challenge by embedding specific design principles and researcher-led practices that uphold analytical rigor while embracing the generative potential of large language models. The following section outlines how reliability, validity, and reflexivity can be preserved, even enhanced, within this dialogic approach.

A core recommendation is to avoid using general-purpose chatbots for serious qualitative analysis, due to risks such as hallucinations and limited transparency about how responses are generated. While some specialized tools attempt to restrict the model's focus to uploaded documents through prompt engineering alone, this approach still carries a higher risk of fabricated content and incomplete traceability. Tools that use retrieval-augmented generation (RAG) architectures offer a more reliable solution. In a RAG system, the model’s outputs are explicitly grounded in specific passages retrieved from the user’s own data, greatly reducing hallucination risk and enhancing transparency about how responses are constructed. Researchers can directly inspect which segments informed each response, enabling validation at the level of source material.

An added benefit of using purpose-built tools is improved data security: processing takes place on designated servers within clearly defined jurisdictions, user data is strictly separated from other users, and it is never used to train the underlying models. These safeguards not only protect sensitive materials but also contribute to the overall trustworthiness and validity of the analytic process. Figure 5 illustrates what this looks like in QInsights.

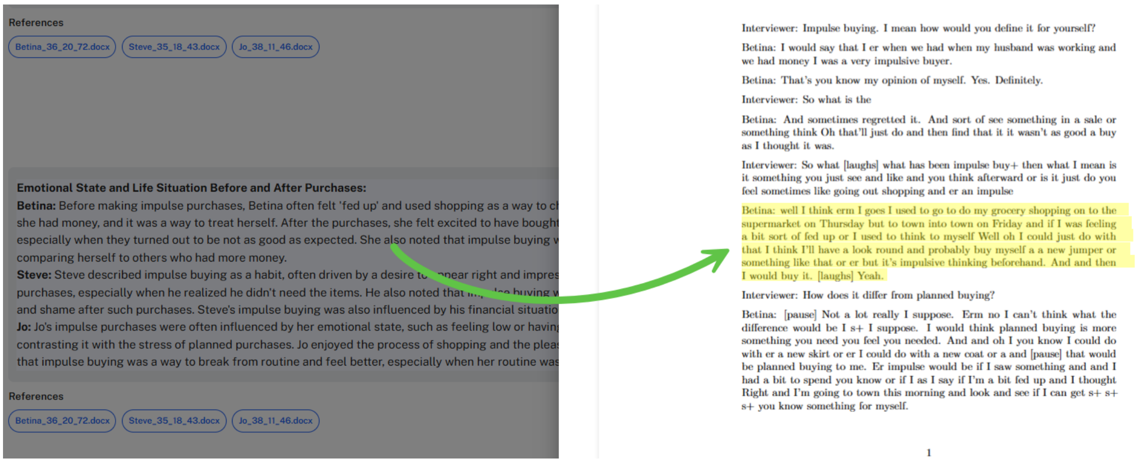

Figure 5: Validation of AI responses within the original source data in QInsights

Since CAAI does not rely on coding, it cannot assess reliability through inter- or intra-coder agreement—a metric whose relevance is itself contested within many interpretive qualitative traditions. Instead, reliability and rigor emerge from the dialogical process between researcher and AI, where question design and synthesis serve as the central units of comparison and reflection. The equivalent of a coding frame is the list of questions developed to explore a topic—this list structures the analytic process and offers a stable basis for assessing consistency.

Moreover, CA^{AI} enhances transparency throughout the analytic process. Because analysis unfolds in dialogue, the entire sequence of questions, AI responses, and researcher reflections is documented in the chat protocol (see Figure 3). This provides a full audit trail of how interpretations were developed, including the reasoning behind each analytic move. In contrast to traditional coding systems—where the final code assignment often obscures the thought process behind it, the dialogic record makes researchers' interpretive pathways visible and reviewable. This added transparency strengthens the validity of findings by allowing for critical scrutiny, replication of questioning strategies, and deeper reflexivity.

transparency yes,
reliability, um, no!

Cross-Person Reliability

To assess reliability across researchers, the same question set can be used by two analysts independently. Both pose the same prompts to the AI assistant, then write their own syntheses based on the AI-generated responses. These syntheses can then be compared and discussed:

- Did the AI return consistent or divergent responses?
- Did both researchers interpret the outputs in similar or contrasting ways?

This dialogic review allows researchers to deepen their findings, uncover blind spots, and, if necessary, revise the question set—rewording unclear questions, adding new ones, or removing irrelevant items. It mirrors the goal of intercoder dialogue but shifts the focus from coding units to interpretive synthesis.

Temporal Reliability

For solo researchers, a reliability check can be performed by repeating the same analytic sequence after a short interval. The researcher reuses the original question set, engages the AI, and writes a second synthesis for comparison. While not a perfect replication strategy, this approach helps assess whether AI responses are stable over time and whether the researcher’s own interpretation remains coherent across analytic cycles.

The temperature setting of the language model plays a critical role here. Lower temperatures reduce randomness, resulting in more predictable, consistent output. When working via API, researchers can set this parameter directly—typically between 0.2 and 0.5 for qualitative research purposes. In contrast, general-purpose chatbots like ChatGPT (which operates at ~0.7) are optimized for balanced creativity and may yield more variation across sessions. This is another reason why dedicated platforms with configurable model settings are preferable for research contexts requiring consistency and traceability.

Table 1: Strategies for Ensuring Quality in CA^{AI}

Quality Criterion	**CA^{AI} Strategy**
Validity / Trustworthiness	Traceability to source data (via RAG); triangulation across groups; iterative question refinement; negative case analysis
Transparency	Full documentation of dialogic interactions via chat protocols; visible interpretive pathways; audit trail of analytic decisions
Reliability / Rigor	Reapplication of refined questions across subgroups; independent synthesis by multiple researchers; replication over time
Reflexivity	Continuous researcher engagement with AI outputs; synthesis written in the researcher's own voice; critical evaluation of emerging insights

In sum, CA^{AI} shifts the focus of quality assurance away from mechanical agreement toward reflexive interpretive rigor. Researchers do not accept AI outputs at face value, but engage with them critically, revisiting the source, questioning assumptions, and refining analytic direction. This not only mitigates the risks of hallucination or misinterpretation, but also aligns with core qualitative values: transparency, reflexivity, and accountability.

While a focused analytic dialogue on a single topic may take only 20–30 minutes, writing the synthesis typically requires more time—ranging from an hour to half a day, depending on the complexity. As the analysis expands across subgroups and topics, depth increases organically. But unlike traditional coding-based methods, where weeks may be spent segmenting text before interpretation begins, every written product in CA^{AI} contributes directly to the final analysis. It's not just faster—it's more conceptually integrated from the start.

In this sense, CA^{AI} does not compromise core qualitative values but reinterprets and operationalizes them for an AI-supported research environment. The emphasis on dialogic engagement, traceability to empirical data, theoretical sensitivity, and reflexive documentation closely aligns with quality principles identified by Strübing et al. (2018). Their call for dynamic, empirically grounded, and theoretically informed processes—rather than rigid proceduralism—finds a contemporary expression within CA^{AI}'s synthesis-driven approach. By embedding AI into an iterative analytic triangle of data, researcher, and assistant, CA^{AI} maintains the integrity of qualitative inquiry while adapting it to new technological possibilities.

What Is Gained and What Might Be Lost by Moving Away from Coding

The move away from coding as the backbone of qualitative analysis, as proposed in CA^{AI}, brings both significant gains and important trade-offs. Recognizing these tensions is essential for ensuring that this paradigm shift remains grounded in the interpretive and ethical commitments of qualitative research.

What is gained is a reorientation of analytic labour—from mechanical segmentation toward conceptual synthesis. Researchers can spend less time organizing data into categories and more time interpreting meaning, engaging directly with patterns, contradictions, and participant voice. The process becomes more flexible, allowing for iterative exploration without being constrained by rigid coding frames. Analysis can begin earlier, proceed more fluidly, and stay closer to the context of the data. In CA^{AI}, researchers can track meaning as it unfolds across documents and subgroups, guided by dynamic questioning rather than static labels. This also opens the door to working with larger and more complex data sets, as analysis is no longer bottlenecked by manual coding capacity.

Another gain is *epistemic clarity*. Instead of collapsing interpretation into category counts, researchers are encouraged to articulate their questions, decisions, and syntheses explicitly. This creates a new form of transparency—where the analytic path can be traced through documented prompts, AI responses, and human reflections. Especially when using retrieval-augmented generation (RAG) systems that anchor outputs in textual evidence, researchers can validate interpretations by returning directly to the source material.

What may be lost, however, is the grounding that traditional coding provides. For many researchers—especially those new to qualitative inquiry—coding offers a clear procedure and a sense of analytic discipline. It externalizes thought and provides a tangible structure for managing complexity. Letting go of coding requires researchers to develop new competencies: asking the right questions, maintaining analytic focus, and resisting the temptation to treat AI outputs as ready-made insights.

Moreover, collaborative analysis may feel less anchored without the familiar scaffolding of coding schemes. In traditional workflows, researchers can code the same dataset, compare overlaps, and calculate inter-coder agreement—providing a clear metric for consistency. In CA^{AI}, this kind of procedural alignment is replaced by interpretive alignment. Researchers collaborate by posing the same questions to the AI assistant and independently crafting syntheses, which are then compared, discussed, and potentially revised. This process lacks the numerical reassurance of agreement coefficients but introduces a more reflexive form of validation—one based on conceptual coherence, critical dialogue, and shared understanding. What is lost in standardization is gained in depth and flexibility, though it demands a greater investment in interpretive transparency.

Finally, certain forms of comparison and quantification such as frequency counts or code co-occurrence matrices play a less central role in dialogic analysis. While numerical metrics are often not essential for meaning-making in qualitative research, they can still be valuable in mixed-methods contexts where empirical traceability and quantifiable outputs are expected. That said, quantification is not entirely absent. For example, identifying which documents a particular theme appears in is both feasible and often valuable. However, relying on LLMs for precise frequency counts can be unreliable at present, especially when working with nuanced or implicit concepts. These limitations are likely to diminish with ongoing improvements in model capabilities, but for now, quantification plays a secondary role in CA^{AI}, supporting rather than driving the analysis.

In sum, moving away from coding is not a rejection of rigor but a repositioning of it. What is gained is interpretive flexibility, depth, and conceptual coherence. What may be lost are procedural anchors and familiar validation tools. The challenge lies in developing new practices that preserve the best of qualitative inquiry while embracing the epistemic possibilities that dialogic AI makes possible.

Paradigm Shifts in Qualitative Analysis: Toward a Post-Coding Era

The analysis method presented in this article, Conversational Analysis with AI (CA^{AI}), is not merely a procedural innovation—it marks a deeper paradigmatic shift in how qualitative research can be conceived in the age of generative AI. While it challenges the long-standing assumption, dominant in coding-based approaches, that systematic segmentation is a necessary precondition for rigorous interpretation, it also resonates with interpretive and reconstructive traditions that have historically emphasized the primacy of meaning-making

you can a 'move away from coding' without using AI over coding. CA^{AI} offers a model in which meaning is co-constructed through structured, iterative dialogue with a large language model.

Other recent approaches signal a move in this direction. Krähnke, Pehl, and Dresing (2025), for example, propose a hybrid method that integrates multiple LLMs into an abductive heuristic process. Their approach treats each model as an independent epistemic actor and constructs meaning through triangulated interaction. It is dialogic, theoretically grounded, and epistemologically reflexive. In this respect, their method stands as a methodological cousin to CA^{AI}. Both approaches emphasize abductive reasoning, iterative engagement, and the researcher's central interpretive role.

Yet while Krähnke et al.'s process involves multiple model perspectives and comparative logic, CA^{AI} operates with a single model and positions the researcher as the sole interpretive agent who orchestrates the analytic dialogue. The emphasis shifts from epistemic triangulation between models to epistemic orchestration through researcher-led questioning. Meaning emerges not through codebook consensus or inter-model dialogue, but through recursive, layered interaction between analyst, data, and model—guided by reflexive judgment and evolving lines of inquiry.

there are endlessly many such options

This reconfiguration marks a paradigm shift. Rather than treating AI as an assistant to accelerate coding or classification, CA^{AI} dispenses with coding altogether. It replaces segmentation and labelling with *questioning and synthesis*. It replaces the logic of breakdown and categorization with a logic of interpretive flow, contextual grounding, and iterative refinement.

Epistemologically, CA^{AI} extends the hermeneutic frame discussed earlier. Knowing is no longer constructed solely through immersion and memoing, but through *conversation*. The AI model acts as a dialogic partner—not because it understands in a human sense, but because its outputs reflect a distributed exposure to language patterns across countless texts, echoing the interwoven lifeworlds researchers aim to explore. As Krähnke et al. argue, LLMs function as abductive catalysts. In CA^{AI}, this potential is fully embraced—but on the condition that the human researcher remains the analytic engine. The model provokes; the human interprets.

This shift has methodological, pedagogical, and ethical consequences. It requires researchers to learn a new craft: not coding but constructing analytic dialogue. It redistributes skill from segmentation to synthesis, from code creation to conceptual orchestration. It alters transparency practices—making question sets, iterative refinements, and synthesis rationales the new basis of analytic traceability. And it calls for new norms around interpretive authority, prompting researchers to continuously reflect on how AI responses are used, rejected, or shaped within their analytic framework.

what is the salient difference between using an AI like this and just a tireless human?

Rather than offering yet another tool to speed up qualitative analysis, *CA to the Power of AI* reframes what analysis is. It proposes that coding is not intrinsic to qualitative rigor, and that a coding-free, dialogic methodology can meet or exceed existing standards of transparency, reflexivity, and depth. It does not simply revise old practices. It reimagines them.

Conclusion

Conversational Analysis with AI (CA^{AI}) marks a turning point in the evolution of qualitative research. Rather than retrofitting generative AI into existing coding-based workflows, it reimagines the analytic process around dialogic engagement—placing interpretation, not categorization, at the centre. This shift reframes the researcher's role from a coder of segments to a conductor of analytic conversation, guiding the AI through structured, reflexive inquiry.

By moving away from segmentation and embracing synthesis, CA^{AI} unlocks new forms of analytical depth, flexibility, and traceability. It preserves the core values of qualitative inquiry while adapting them to a technological landscape shaped by distributed intelligence and iterative interaction. CA^{AI} does not dilute rigor; it redistributes it. And in doing so, it challenges longstanding assumptions about what constitutes qualitative analysis and how knowledge is constructed.

Leveraging the strengths of large language models for data retrieval and synthesis, CA^{AI} keeps the researcher firmly in control—through structured questioning, interpretive judgment, and iterative refinement. This approach offers not only gains in efficiency and flexibility but also invites a fundamental reconsideration of epistemological assumptions and validation practices within qualitative inquiry.

For those willing to step beyond coding and embrace AI as a collaborative sense-making partner, CA^{AI} offers a compelling alternative. Its promise lies in its capacity to deepen interpretation, democratize analytic access, and expand the epistemic horizon of qualitative research in the age of artificial intelligence.

References

Charmaz, K. (2024). Constructing grounded theory (3rd edition). SAGE Publications.

Deiner, E., Goldberg, M., McDuff, D., & Acquisti, A. (2024). *Thematic analysis of a social media corpus with a single prompt*. [Manuscript submitted for publication].

Corbin, J., & Strauss, A. (2014). Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory (4th ed.). SAGE Publications

Friese, S. (2023a). Rethinking Qualitative Data Analysis: Do we truly want a faster horse? Qeludra Blog. https://qeludra.com/blog/rethinking-qualitative-data-analysis-do-we-truly-want-a-faster-horse

Friese, S. (2023b). Life without coding: *Does Generative AI Fulfill This Dream?* Qeludra Blog. https://qeludra.com/blog/life-without-coding-qualitative-analysis

Friese, S. (2024). *Ethical and Responsible Use of AI for Qualitative Analysis.* Qeludra Blog. https://qeludra.com/blog/ethical-aspects-of-ai-in-research-part2

Friese, S. (2025). Generative AI: A Catalyst for Paradigmatic Change in Qualitative Data Analysis. In Christou A. Prokopis (Ed.) *Artificial Intelligence (AI) in Social Research*, Chapter 7. CAB International, Wallingford, UK.

Gamieldien, Y., Case, J., & Katz, A. (2023). Advancing qualitative analysis: an exploration of the potential of generative AI and NLP in thematic coding. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4487768

Gao, J., Choo, K. T. W., Cao, J., Lee, R. K., & Perrault, S. T. (2023). CoAIcoder: examining the effectiveness of ai-assisted human-to-human collaboration in qualitative analysis. *ACM Transactions on Computer-Human Interaction*, 31(1), 1-38. https://doi.org/10.1145/3617362

Goyanes, Manuel & Lopezosa, Carlos & Jordá, Beatriz. (2024). *Thematic Analysis of Interview Data with ChatGPT: Designing and Testing a Reliable Research Protocol for Qualitative Research.* SocArXiv. https://doi.org/10.31235/osf.io/8mr2f_v1

Hayes, A. (2025). "Conversing" with qualitative data: enhancing qualitative research through large language models (LLMs). International Journal of Qualitative Methods, 24. https://doi.org/10.1177/16094069251322346

Hitch, D. (2024). Artificial intelligence augmented qualitative analysis: The way of the future? *Qualitative Health Research, 34*(7), 595–606. https://doi.org/10.1177/10497323231217392

Ibrahim, M., & Voyer, D. (2024). The augmented qualitative researcher: Using generative AI in qualitative research. [Manuscript submitted for publication].

Krähnke, U., Pehl, T., & Dresing, T. (2025). Hybride Interpretation textbasierter Daten mit dialogisch integrierten LLMs: Zur Nutzung generativer KI in der qualitativen Forschung. SSOAR. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-99389-7

Lixandru, I. (2024). The use of artificial intelligence for qualitative data analysis: chatgpt. Informatica Economica, 28(1/2024), 57-67. https://doi.org/10.24818/issn14531305/28.1.2024.05

Mannheim, K. (1936). Ideology and utopia: An introduction to the sociology of knowledge (L. Wirth & E. Shils, Trans.). Harcourt, Brace & Company. (Original work published 1929)

Morgan, D. (2025). Query-Based Analysis: A strategy for analyzing qualitative data using ChatGPT. Qualitative Health Research, forthcoming.

Nguyen-Trung, K., & Nguyen, N. L. (2025, March 4). Narrative-Integrated Thematic Analysis (NITA): AI-Supported Theme Generation Without Coding. SocArXiv. https://doi.org/10.31219/osf.io/7zs9c_v1

Perkins, M., & Roe, J. (2024). The use of generative AI in qualitative analysis: Inductive thematic analysis with ChatGPT. *Journal of Applied Learning & Teaching, 7*(1), 390–405. https://doi.org/10.37074/jalt.2024.7.1.22

Schreier, M. (2012). *Qualitative content analysis in practice*. SAGE Publications.

Strübing, Jörg, Hirschauer, Stefan, Ayaß, Ruth, Krähnke, Uwe and Scheffer, Thomas. "Gütekriterien qualitativer Sozialforschung. Ein Diskussionsanstoß" *Zeitschrift für Soziologie*, vol. 47, no. 2, 2018, pp. 83-100. https://doi.org/10.1515/zfsoz-2018-1006

Thominet, L., Amorim, J., Acosta, K., & Sohan, V. K. (2024). Role play: conversational roles as a framework for reflexive practice in ai-assisted qualitative research. Journal of Technical Writing and Communication, 54(4), 396-418. https://doi.org/10.1177/00472816241260044

Xiao, Z., Yuan, X., Liao, Q. V., Abdelghani, R., & Oudeyer, P. (2023). Supporting qualitative analysis with large language models: combining codebook with gpt-3 for deductive coding. 28th International Conference on Intelligent User Interfaces, 75-78. https://doi.org/10.1145/3581754.3584136

Zackary, O. D. (2024) Scalable qualitative coding with LLMs: Chain-of-thought reasoning matches human performance in some hermeneutic tasks. arXiv. https://arxiv.org/abs/2401.15170

Zhang, H., Wu, C., Xie, J., Rubino, F., Graver, S., Kim, C., Carroll, J.M., & Cai, J. (2024). When Qualitative Research Meets Large Language Model: Exploring the Potential of QualiGPT as a Tool for Qualitative Coding. ArXiv, abs/2407.14925. https://doi.org/10.48550/arXiv.2407.14925