Generative AI: A Catalyst for Paradigmatic Change in Qualitative Data Analysis

Susanne Friese

Abstract

The aim of this chapter is to investigate whether Thomas Kuhn's concept of scientific revolutions, as described in his seminal work *The Structure of Scientific Revolutions* (1962), can be extended to examine shifts in qualitative research methodologies driven by the advent of artificial intelligence (AI). Initially, Kuhn's process of scientific revolution will be outlined, focusing primarily on the dynamics detailed in his landmark text. While Kuhn's discussion was predominantly confined to the natural sciences, there is ongoing debate about its applicability to the social sciences. This debate will be addressed, arguing for the framework's relevance beyond the natural sciences, as Kuhn himself suggested in a postscript to a later edition (2012). Building on this foundation, the history of one of the predominant methodologies in qualitative data analysis, coding, will be explored. The exploration will consider how the emergence of AI technology could potentially induce a Kuhnian-type shift, leading to significant transformations in how qualitative data will be analysed in the future.

Cite as:

Friese, S. (2025). Generative ai: a catalyst for paradigmatic change in qualitative data analysis. *Artificial Intelligence (AI) in Social Research*, 71-83. https://doi.org/10.1079/9781800626607.0007

7.1 Introduction

This chapter explores the transformative potential of AI on qualitative data analysis approaches, applying the framework of Thomas Kuhn's theory of scientific revolutions (Kuhn, 1962). The chapter starts by outlining Kuhn's ideas before shifting to analysing how and why these ideas can be applied to the social sciences and especially to the effect the transformative new technology has on research methodology. It critically examines how artificial intelligence (AI) is reshaping traditional methodologies and prompting a re-evaluation of established analytical frameworks. The historical development of coding as a central technique in qualitative analysis will be traced, emphasizing its role in structuring and systematizing the research process. The recent integration of AI into coding practices will then be outlined, highlighting both the opportunities and challenges presented by such technological advancements. This discussion is framed within the context of Thomas Kuhn's theory of scientific revolutions, drawing parallels between the disruptive influence of AI and the paradigm shifts described by Kuhn. As anomalies and inefficiencies in traditional coding practices are magnified by the capabilities of AI, a state of crisis is emerging within the qualitative research community. This growing crisis, underscored by a clash between traditional methodologies and innovative AI-driven approaches, along with the community's resistance to new technologies, invites us to explore these dynamics through the lens of Kuhn's theory of scientific revolutions. Let us first examine the distinct phases of a paradigm shift as described by Kuhn to better understand how such shifts might unfold in the context of qualitative data analysis.

7.2 Scientific Revolutions Through Kuhn's Lens

7.2.1 The phase of normal science

Kuhn argued that scientific progress does not occur in a smooth, linear trajectory. Instead, he proposed a model characterized by alternating periods of 'normal science' and 'revolutionary science'. During phases of normal science, scientists largely agree upon a dominant paradigm (Lakatos and Musgrave, 1970; Kuhn, 2012). This paradigm represents a shared framework of established theories, widely accepted methods and a common set of underlying philosophical assumptions. It acts as a lens through which researchers perceive the problems they address and the validity of potential solutions.

Within the boundaries of a normal science period, researchers engage in what Kuhn calls 'puzzle-solving'. They do not challenge the fundamental assumptions of the paradigm itself. Instead, they work to refine existing theories, extend them to explain

observed anomalies and explore new applications for established knowledge. It can be thought of as filling in missing pieces within an agreed-upon picture.

This commitment to a dominant paradigm allows focused and efficient progress. Scientists build upon a shared language and set of tools, trusting in the ability of the current framework to ultimately address any challenges that arise. However, Kuhn (2012) and Feyerabend (1975) argued that this intense focus on a single paradigm can lead to an over-reliance on established frameworks, potentially blinding researchers to phenomena or solutions that exist outside its prescribed boundaries. This tunnel vision can ultimately precipitate the kind of revolutionary changes that Kuhn describes.

7.2.2 Occurrence of anomalies

Anomalies play a critical role in this process. They are puzzling observations or experimental results that stubbornly resist explanation within the existing scientific framework. While these anomalies are disruptive, they do not immediately indicate the failure of a paradigm. Scientists often exhibit initial resistance to discarding their well-established theories, preferring to dismiss anomalies as experimental errors, isolated incidents or simply mysteries that will eventually be resolved by the current paradigm. This tendency to adhere to and operate within the existing framework underscores both the efficiency of using a shared paradigm and a steadfast belief in its capacity to solve problems.

Kuhn asserts that anomalies hold transformative potential. A single anomaly might be dismissed, but when anomalies persist or a cluster of related irregularities arise, cracks begin to appear in the seemingly solid foundation of the paradigm. Scientists can no longer simply relegate problems to the margins; they are forced to question whether their fundamental assumptions are somehow flawed or incomplete. It is this process, fuelled by nagging anomalies that refuse to be tamed, that can ultimately pave the way towards a paradigm shift and revolutionize scientific understanding.

As anomalies accumulate and stubbornly defy resolution within the boundaries of the existing paradigm, a sense of unease begins to permeate the scientific community. The once-unshakeable confidence in the dominant framework falters as scientists confront the limitations of their current understanding.

7.2.3 State of crisis

In Kuhn's theory, a 'crisis' in science happens when existing theories cannot explain new or persistent problems. During such crises, scientists start looking for solutions more broadly. They consider new and unusual ideas that they might have ignored before. This is different from normal times when scientists usually stick to well-established theories. A crisis shakes up the old ways of thinking and can lead to major changes in scientific understanding.

A good example of this is the shift from Newtonian physics to Einstein's theory of relativity. For many years, Newton's laws were seen as the ultimate explanation for the forces of nature. However, as scientists began to explore the universe more deeply, they encountered observations that Newtonian physics could not explain, such as issues with the orbit of Mercury and the speed of light. This led to a scientific crisis.

As the crisis deepened, Albert Einstein introduced his theory of relativity, which provided solutions to these problems and offered a new way of understanding gravity and motion. Einstein's theory was initially seen as radical and faced scepticism because it challenged the well-established Newtonian framework. However, its ability to solve problems that were inexplicable under the old paradigm led to its acceptance and a revolutionary shift in physics.

This transformative change in scientific thinking frequently involves an overhaul of foundational concepts, methodologies and even the types of research questions considered legitimate. The landscape of the discipline is irrevocably altered. Thus, a crisis acts as a prelude to the birth of a new paradigm.

7.2.4 Birth of a new paradigm

Kuhn underscored the radical nature of this change by introducing the concept of 'incommensurability' (c.f. Feyerabend, 1975). Competing paradigms, he posited, may be so fundamentally different in their assumptions and conceptual language that direct comparison or translation becomes impossible. It is as if scientists operating within different paradigms are speaking entirely separate dialects, unable to fully understand each other's perspectives. Therefore, a true paradigm shift is not a process of gradual evolution but a revolution, a decisive break from the past. The lens through which scientists observe their field of study changes so drastically that the old ways of thinking and operating become obsolete. This revolution gives birth to a new form of 'normal science', operating under the rules and assumptions of the freshly established paradigm, until the cycle – anomalies, crisis, revolution – eventually begins again.

Kuhn underscored the profound and often unsettling nature of a paradigm shift. It is not a matter of mere refinement or adjustment but involves embracing a new way of seeing and thinking about the world. This change can be met with resistance, as established scientists may have deeply internalized the old worldview, making the transition to a radically different framework both intellectually and psychologically challenging. Embracing a new paradigm, according to Kuhn (2012), is rarely driven by purely logical deduction or hard empirical evidence alone. While the accumulation of anomalies that resist explanation within the old paradigm plays a significant role, the shift is ultimately rooted in a complex interplay of factors. One crucial element is the emerging paradigm's potential to solve the problems that plagued its predecessor. A new paradigm must not only acknowledge the old paradigm's failings but also offer a compelling promise of

illuminating previously unresolved puzzles. Its ability to provide novel explanations or address lingering anomalies holds great persuasive power within the scientific community.

Additionally, the decision to embrace a new paradigm is influenced by its potential to open new avenues of research. A promising paradigm generates a sense of excitement, suggesting fertile ground for exploration, novel questions and the potential for significant discoveries. This forward-looking allure can be a deciding factor, especially for younger scientists less wedded to the old ways of thinking.

Finally, Kuhn stressed that the resolution of revolutions cannot be separated from the social dynamics of the scientific community. Arguments for and against competing paradigms involve elements of persuasion, reputation and generational shifts. The decision to embrace a paradigm shift is never a purely objective one; it is influenced by the power structures, personalities and intellectual currents within a particular scientific field.

In essence, the abandonment of an old research tradition in favour of a revolutionary new paradigm is the result of a complex convergence of factors. It is driven by the limitations of the old paradigm, and the problem-solving potential and promise of the new, as well as the ever-evolving social and intellectual landscape of the scientific community itself.

7.3 Application of Kuhn's Concept of Paradigm Shift to Other Fields

Kuhn focused primarily on the natural sciences in his seminal work. Critics have contended that the processes he described are less applicable to the social sciences, where there is less emphasis on discovering universal laws and more on interpreting human behaviour and social structures (Bird, 2022). Furthermore, the social sciences are distinguished by a rich diversity of coexisting theories and methodologies, all of which can be equally valid. This methodological pluralism reflects the varied approaches used to understand human behaviours and societies, indicating that multiple paradigms often operate concurrently without one achieving the dominance necessary to marginalize the others. Common paradigms in social sciences include positivism, interpretivism, Marxism, feminism, critical theory and post-modernism, each providing a distinct perspective on the social world. This contrasts sharply with Kuhn's concept of a single dominant paradigm that governs the practice of normal science, where a new paradigm, like Einsteinian physics, replaces an old one, such as Newtonian physics.

This argument, however, does not reflect the motivation of why Kuhn wrote the book. He wanted to challenge the prevailing view of science as a purely objective and linear progression of discoveries. Kuhn argued that scientific progress is a fundamentally human endeavour and not merely the result of a methodical and rational process (Kálmán, 2016). In the opening lines of his book, Kuhn stated that he aimed to demonstrate how textbooks, the traditional mediums through which each new generation of scientists learns its craft, have fundamentally misled us. Throughout the book, he deconstructed this view chapter by chapter, describing how shared beliefs within a community are formed, how they sustain that community over time and how they are eventually disrupted by anomalies, leading to a crisis and transformative shift in these shared beliefs and practices. And this process, the author would argue, can be transferred to what is currently happening to qualitative research methodologies, and particular data analysis.

7.4 How Emergent Technologies Reshape the Practice of Coding in Qualitative Data Analysis

Coding has become a central technique in many forms of qualitative research (Braun and Clarke, 2006, 2023; Bazeley, 2013; Charmaz, 2014; Kuckartz, 2014; Silver and Lewins, 2014; Corbin and Strauss, 2015; Friese, 2019; Miles *et al.*, 2019; Richards, 2020; Saldaña, 2021; Mayring, 2022). Some of the books listed above have been published in ten or more editions. The widespread adoption of coding across multiple qualitative research methods and its incorporation into software designed specifically for qualitative analysis has standardized certain analytical processes and promoted a consistent methodological approach within the field. However, it is important to note that, while coding is influential and widely adopted, it does not strictly conform to Kuhn's definition of a paradigm. It lacks a singular theoretical commitment, is highly adaptable across various theoretical frameworks, and is used by a diverse community of researchers without a unified set of beliefs about its function and implications.

Therefore, while coding exemplifies more of a methodological consensus rather than a full paradigm, it serves as a useful lens to examine how emerging technologies could reshape established practices in qualitative research. The advancements in AI, particularly the arrival of ChatGPT, have introduced a significant shift in the research landscape. Many qualitative researchers find themselves grappling with this new technology, marked by a proliferation of unanswered questions, intense debates between early adopters and critics, and a general lack of understanding. This unsettled environment is precipitating what could be described as a state of crisis, reminiscent of the conditions that Kuhn outlines as preceding a paradigm shift.

The current dynamics within the community of qualitative researchers echo the transformative shifts Kuhn described in his model of scientific revolutions. As coding has been integral to data analysis in qualitative research, the integration of AI challenges and potentially redefines this established methodological approach. Researchers are now confronted with the possibility that AI can automate or augment the coding process (Hitch, 2023), or even fundamentally alter the analysis process, raising both opportunities and concerns that resonate with Kuhn's description of a scientific community in transition. Just as anomalies within a paradigm lead to questions that challenge established norms, the capabilities, and limitations of AI in qualitative research are prompting a re-evaluation of long-held assumptions about what constitutes rigorous and insightful analysis.

This intersection of coding as an established practice and the disruptive influence of AI technology suggests that we may be on the cusp of a methodological revolution, in qualitative research. The friction between traditional methodologies and innovative technologies illustrates a classic Kuhnian crisis, setting the stage for a significant transformation in how qualitative research is conducted. This transformation will lead to the establishment of new standards, practices and understandings that better accommodate the capabilities and challenges of AI.

7.4.1 Period of normal science: how coding became a prevalent approach for qualitative data analysis

The term 'coding' in qualitative research did not emerge as a planned terminological invention but as an evolution in the methodology of data analysis. In early method textbooks, the need for systematic approaches to organize and analyse data was already recognized. For instance, Palmer (1928) noted that 'skilled research workers are able to rearrange the data observed in interviews so that remarks concerning one topic can be grouped together without losing the implication contained in the original sequence'. This early insight highlights the foundational need for what would later become known as coding.

By 1937, the term 'coding' was used explicitly within the marketing research community, as documented by the Association of American Marketing, to describe the process of classifying open-ended survey responses. This marked one of the earliest formal adoptions of the term in a research context. Despite these early uses, the term 'coding' did not gain widespread recognition immediately. Influential researchers such as Robert K. Merton developed elaborate systems to analyse unstructured interviews but did not use the term 'coding'. Similarly, Adorno *et al*. (1950) referred to processes akin to coding as 'categorization' or 'classification' in their reports. It was not until the mid-1950s that the terminology began to catch on, as noted by Hyman and Sheatsley (1954), who frequently employed the terms 'code', 'codes' and 'coding'.

The term's first sustained use is often associated with the methodological writings that emerged from the study 'Boys in White' by Becker *et al*. (1961). This research, focused on higher education, a field then dominated by psychology, required explicit descriptions of the researchers' procedures and logic. This need probably catalysed the formal adoption and regular use of 'coding' in their reports. Around the same time, Whyte's (1955) book entitled *Street Corner Society* marked a significant shift from case-based to code-based analysis, influenced by anthropological approaches such as those of Bronisław Malinowski. This shift was further propelled by technical advancements such as the portable tape recorder, which allowed researchers to capture real-time data and necessitated more refined categories of analysis. As Chapoulie (1987) noted, the precise verbatim accounts from field notes demanded that researchers construct finer categories of analysis and provide detailed explanations of their interpretations.

Before the advent of computer-assisted qualitative data analysis (CAQDAS) software, these developments set the stage for coding to become an essential element of qualitative research. The systematic procedures used to extract relevant material from field notes, such as marking up field notes with different coloured pencils to indicate different topics, laid the groundwork for what would later be facilitated by software. Becker *et al.* (1984) described these early methods as 'loose inclusive' coding procedures, where codes were assigned whenever there was any indication of relevance to the current issue, highlighting a systematic approach to uncovering all relevant items to a particular proposition.

Building on the historical narrative of the development of the term 'coding' in qualitative research, Fielding and Lee (1998) noted that while Becker and Geer (1960) were instrumental in introducing the term, it was Bernard Glaser and Anselm Strauss (1967) who popularized it through their seminal work, *The Discovery of Grounded Theory: Strategies for Qualitative Research*. In this book, they frequently used the term 'coding', applying it extensively as part of their methodological framework without specific elaboration on its usage.

The widespread adoption and popularity of grounded theory, especially through the later works of Strauss and Corbin, such as *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (1998), cemented 'coding' as a fundamental term and procedure in qualitative data analysis. As qualitative research evolved with technological advancements, the development of software to support data analysis incorporated these established methodologies. Early qualitative data analysis software varied in terminology, with some initially using terms such as 'indexing' or 'nodes' to describe processes akin to coding. However, as the practice standardized, the terms 'coding' and 'codes' became universally adopted, reflecting the established terminology within the research community. This shift not only standardized the

language used across different platforms but also reinforced the integral role of coding in qualitative analysis.

Building on the historical context of the term 'coding' in qualitative research and its methodological evolution, the introduction of CAQDAS software in the mid-1980s represented a significant technological leap that transformed the landscape of data analysis. Richards and Richards (1991) described this transformation as a methodological revolution, noting that 'every aspect of the handling of qualitative data is dramatically changed by computers'. Below it will be argued that it was more an evolution rather than a revolution. None the less, it had a significant impact on how researchers analysed qualitative data. CAQDAS software enabled a myriad of new capabilities: it facilitated the segmenting of text to finely defined excerpts or to information about the text, such as characteristics of respondents or aspects of documents. This shift allowed a more nuanced interplay between the original textual data and the conceptual frameworks that researchers use and create for exploring their data. Information could now be organized in a tree-structured category system or a code list, serving as a repository for document references and reflective thinking.

Furthermore, CAQDAS software enhanced the efficiency and depth of qualitative analysis by allowing multiple indexing of text segments, searching for patterns and overlaps in coded units, and linking to detailed information about documents or respondent characteristics. Researchers could save query results automatically for further analysis, offering a set of logical analysis operations that were quick, secure and capable of frequent revisions. The creation of new categories, the modification of relationships between categories and changes to the category system could all be managed dynamically, allowing quasi-statistical analyses, summary tables and the linking of disparate data. Conclusions could be checked and verified with greater rigour, challenging the traditional dichotomy between quantitative and qualitative methods, and promoting methodological diversity in qualitative research.

CAQDAS software has significantly impacted the operational aspects of coding and data analysis, enhancing the traditional processes. While fundamentally researchers continued to engage in coding, they did so in a more refined and sophisticated manner. The technology enabled a dynamic interplay between text and encoded data, allowing researchers to seamlessly toggle between raw data and their analytical constructs. While CAQDAS software brought about significant methodological changes and reshaped the research process, its introduction was more of an evolutionary development in the capabilities of qualitative researchers rather than a fundamental shift. It augmented and expanded existing methodologies, allowing deeper, more systematic exploration of complex qualitative data.

Despite these advancements, the adoption of CAQDAS software also came with notable reservations. Initial concerns included a preoccupation with quantity over depth and

form over content, suggesting that the technology might lead researchers to prioritize extensive data handling at the expense of in-depth analysis. Additionally, there was apprehension that such technology could distance researchers from their data, reducing the intimate familiarity that qualitative research traditionally requires. Furthermore, researchers were concerned that not everyone would be able to use the technology due to high costs and a steep learning curve (Fielding and Lee, 1998).

CAQDAS software has not only revolutionized the operational aspects of qualitative research but has also cemented coding as a widely accepted method of data analysis. Over the years, it has become integral to the qualitative research process, underscoring the value of systematic, structured analysis in deriving meaningful insights from complex datasets. As CAQDAS platforms evolved, they facilitated increasingly sophisticated methodologies, making coding indispensable across various disciplines and studies. Using Kuhn's framework, this is comparable with the 'normal science' period.

Currently, this established method faces a new challenge: the advent of AI technology. Researchers are currently experimenting with the new technology, navigating the implications of integrating AI with traditional coding practices.

7.4.2 Anomalies: attempts to integrate AI technology into current coding practices

The coding approach, with its emphasis on categorizing qualitative data into discrete codes, has served the field of qualitative research faithfully for decades. However, with the rise of AI, cracks are beginning to appear. Shortly after the release of ChatGPT, the first CAQDAS programs began incorporating AI technology in various ways. One notable implementation was allowing AI to code the data. This process involves the AI scanning through the data paragraph by paragraph, assigning multiple codes to each section. At first glance, and when examining individual paragraphs, the results might appear satisfactory. However, AI lacks the capability to perform a critical function that human coders routinely undertake – comparing and contrasting the content across different paragraphs.

Human coders often evaluate whether a code needs to be renamed to better encompass the content from another paragraph or decide if an existing code should be applied instead. In contrast, AI tends to generate mostly new codes for each paragraph without such comparative analysis. Consequently, this approach can lead to a proliferation of codes, with the number growing exponentially as more documents are processed as shown in Table 7.1.

Table 7.1. Results of AI-coded documents.

Number of documents	Number of codes
Two documents, 27 pages	393
Four 1-hour interview transcripts	577
16 documents (literature)	2807
37 documents (company sample project containing 30 two-page interviews, seven articles)	2347

The task of consolidating the AI-generated codes remains the responsibility of the researcher. In testing this feature, the author discovered that aggregating the AI-generated codes often takes longer than coding the data manually from scratch. Perhaps due to this inefficiency, another CAQDAS tool opts not to offer automatic coding. Instead, it provides options to suggest codes for highlighted data segments or to suggest subcodes for a higher-level code that already includes several segments manually coded by the user. By examining these features during numerous workshops, researchers found them to be helpful, akin to receiving a second opinion when no collaborative partner is available to discuss the data. However, while these tools are appreciated for their support, they do not necessarily enhance the efficiency of the analysis process. On the contrary, they may even increase the time required due to the need for additional oversight and adjustment.

Common features now implemented in CAQDAS tools include options to use AI for summarizing documents or coded data segments. While summaries of coded data segments tend to be more accurate due to their focus on a specific topic, they still require verification by the researcher. A fundamental issue with this process is the lack of direct input from the user. The user cannot specify details about their project, such as the nature of the data sources (e.g. interviews or focus group transcripts), nor can they request that the AI prioritize certain themes such as issues a, b and c in its summaries. This level of specificity is likely to lead to more relevant and accurate summaries. At present, the AI-assisted tool often relies on educated guesses. For instance, it may correctly identify data as coming from focus group transcripts in one instance but incorrectly assume they are from a single interview respondent in another.

In response to the critiques, software developers have attempted to enhance the functionality of CAQDAS tools by integrating additional features. For example, they now allow users to input context information for a project to refine the automated coding process, or ask the AI assistant to code a document based on a code that is defined by the user. Additionally, some tools offer the option for users to engage in a chat with documents.

Despite these enhancements, it essentially circles back to the fundamental requirement of coding. Users must still code the data to leverage most of the software's analytical capabilities. Users must **still code** the data to leverage most of the software's analytical capabilities.

These implementations illustrate a critical misalignment: AI capabilities are being constrained within an established framework that may not fully exploit their potential. The current approach presses AI to operate within the confines of coding-centric methodologies, which may not be the most effective use of what AI can offer. These limitations, however, are not confined solely to software development. Many qualitative researchers also struggle to move beyond the methodological approaches with which they are familiar. Various research groups and academics have experimented with using ChatGPT for coding and thematic analysis (Adduesselam, 2023; Gao *et al.*, 2023; Hitch, 2023; Morgan, 2023; Smen *et al.*, 2023; Xiao *et al.*, 2023; Zhang *et al.*, 2023; Christou, 2024; Goyanes *et al.*, 2024; Nguyen-Trung, 2024). While most authors are enthusiastic about the potential that the new technology offers, attempts to apply traditional coding methods using this technology have revealed significant shortcomings. Below are some of their conclusions.

Nguyen-Trung (2024) documented an attempt to utilize GPT-4 for coding an entire dataset. The results were unsatisfactory; the AI's output was overly concise, with each code encapsulating only a single interview quote. This approach risked omitting substantial information from other interviews, demonstrating that despite its advanced capabilities, GPT-4 struggled to manage and synthesize large sets of qualitative data comprehensively. The coding produced was too narrow, failing to capture the broader thematic patterns necessary for robust qualitative analysis.

Similarly, Goyanes *et al.* (2024, p. 25) criticized the quality and granularity of thematic patterns produced by these platforms, noting that the subtleties and contextual insights crucial to qualitative research were lacking: 'Our examination of the platform in terms of the quality and granularity of thematic patterns is rather unsatisfactory when it comes to the subtle nuances and contextual insights generally associated with qualitative research.'

Furthermore, Gamieldien *et al.* (2023) described issues with inconsistent coding by a natural language processing (NLP) GPT model, which often mislabels thematic clusters or repeats code names inconsistently: 'Turning to the NLPGPT method, the model labelled each cluster with different names even when it spoke about the same idea and only repeated the same names of codes on a handful of them.' These experiences echo the problem explained above where AI generates an excessive number of codes without effective aggregation.

Another significant limitation that is often overlooked concerns the restricted context window of current large language models (LLMs). While there is optimism that AI could

enable researchers to analyse vast quantities of data swiftly (e.g. Jiang *et al.*, 2021; Theelen *et al.*, 2024), practical applications have been limited. Most experiments have involved only one or two documents. Moreover, attempts to analyse larger sets, such as ten or more interviews simultaneously, have proven either unfeasible or have yielded superficial results (Nguyen-Trung, 2024). This issue arises because once the volume of data surpasses the LLM's context limit, the model tends to rely on summaries rather than the original text, leading to a potential loss of depth and detail in the analysis.

These challenges represent significant deviations from expected outcomes. Just as anomalies in Kuhn's narrative signal the limitations of a prevailing scientific paradigm, these issues highlight the constraints of applying AI within the current framework of coding.

Given these findings, there is a compelling argument to be made for re-evaluating the integration of AI in qualitative data analysis. Instead of forcing AI to adapt to the traditional coding paradigm, there may be a need to rethink and possibly redesign AI tools to better align with the objectives and demands of qualitative research (D.L. Morgan, unpublished data).

This view is not shared by anyone. Kuckartz and Rädiger (2024) argue that generative AI will not be disruptive. Instead it can be integrated into established analytical methods. Moreover for AI to be implemented effectively, a proven methodological framework is required.

7.4.3 Resistance and crisis within the qualitative research community

The resistance encountered during these integration attempts illustrates a scenario reminiscent of a Kuhnian paradigm shift. Researchers who are deeply entrenched in established methods of data analysis often view AI as a potential threat rather than a revolutionary opportunity. This resistance became particularly apparent during spirited debates on the Qualitative Social Research mailing list (QSF Mailing List, n.d.) of which the author is a member and was further evidenced by several universities restricting access to AI-assisted functions in CAQDAS software, citing privacy concerns without substantial legal justification.

Addressing data privacy and protecting respondents are undoubtedly important, yet there must also be a willingness to explore new technologies rather than dismissing them outright. It is a common misconception that the providers of these tools will use the data to train their models. However, most platforms offer users the option to opt out of data sharing for training purposes in their settings. Furthermore, when LLMs are accessed via an application programming interface (API) within a CAQDAS tool, the data are never used for training.

Other concerns include an over-reliance on technology and potential erosion of the qualitative researcher's interpretive authority. Many in the field are hesitant to embrace a perspective where, as one researcher described it, 'meaningful and creativity-based material interpretation becomes an engineering problem to be solved'. Despite this resistance, there is recognition that AI is an enduring presence, creating tension between the comfort of familiar methods and the uncertain potential offered by technological advancements.

As coding is the basis for widely used approaches to qualitative research, a broad range of researchers have a stake in maintaining coding as essential to qualitative research itself (D.L. Morgan, unpublished data). Researchers struggle to envision how AI can be successfully integrated into their established workflows. This juxtaposition highlights the challenges of navigating through a period of significant technological transformation in qualitative research.

7.4.4 The need for a new perspective: a conversational approach to qualitative data analysis

Transitioning to a new perspective in qualitative data analysis is crucial for maximizing AI's potential in the field. The key is not to force AI into the confines of the coding approach but rather to imagine entirely new ways of working with data, where AI acts as a collaborator and an assistant to the researcher, freeing them from the constraints of manual coding. Embracing AI as a partner, rather than as a replacement, offers a pathway towards re-imagining how we conduct qualitative data analysis. The crisis is an opportunity – an impetus for re-evaluation and the potential for ground-breaking new methodologies that leverage the best of both human intuition and the power of AI.

In a series of webinars and presentations, the author has introduced 'Conversational Analysis with AI' as a new method that involves directly engaging with an AI assistant in a dialogue format. This approach does not seek to automate the analysis process, nor does it rely on the AI assistant to independently identify all 'relevant' themes. Instead, the analysis begins with the researcher posing broad, exploratory questions about specific aspects of their data.

Following the initial response from the AI, the researcher delves deeper by asking more targeted follow-up questions. Often, something in the AI's response will catch the researcher's attention, prompting further inquiry. Depending on the nature of the project, the researcher might start by focusing on one document at a time, gradually incorporating more documents into the analysis as the investigation deepens.

Once a topic has been thoroughly explored, the researcher documents their findings and asks the AI assistant to retrieve supporting quotes. The researcher then reviews the first write-up with the AI, evaluating whether the interpretations align with the data,

checking for any oversights and ensuring the logical consistency of the arguments before moving on to the next topic.

The next topic might be predetermined or could emerge from the analysis of previous data. The analyst applies the same iterative process as before. This often leads to the discovery of relationships within the data as the researcher progresses through subsequent topics. With a comprehensive understanding of the data from the ongoing discussion, the AI assistant can effectively verify or refute the interpretations.

This collaborative, conversational approach allows a dynamic and interactive exploration of the data, leveraging the AI's capabilities to enhance and expand the analytical process. Through this method of continuously engaging in a constructive dialogue, the analysis gains depth and breadth.

D.L. Morgan (unpublished data) suggested a similar approach termed 'query-based analysis', detailed in three progressive steps. The first step consists of posing broad, undirected questions to unearth foundational concepts within the data, laying the groundwork for a detailed analysis. In the second step, researchers delve deeper by asking more specific questions to expand on the initial themes identified. The third and final step involves a thorough examination of the text to extract and select supporting quotations that will be included in the research report.

7.4.5 The transformation

D.L. Morgan (unpublished data) posited that AI technology could significantly transform qualitative research by potentially eliminating traditional coding processes and consequently reducing the reliance on specialized CAQDAS software. This development could lower the barriers to analysing qualitative data, as there is typically a steep learning curve associated with mastering these sophisticated tools. Making the process more accessible could extend its use beyond traditional researchers to a broader audience. To facilitate this vision of qualitative data analysis, the author has developed the AI-powered software QInsights (further detail available at: https://qeludra.com, accessed 2 October 2024).

It has also been speculated that the extent to which this new technology will disrupt the field of qualitative research depends largely on various social factors (D.L. Morgan, unpublished data). A critical determinant will be the reaction from stakeholders deeply entrenched in the traditional, coding-centric methods of qualitative analysis, such as academic institutions, journal editors and funding bodies. These groups often serve as gatekeepers within the academic community, and their acceptance or resistance to AI-driven methods will significantly influence how these technologies are adopted and their subsequent impact on the field.

It may take some time before we witness a shift in the standard approaches to qualitative data analysis. As discussed earlier, this change does not encompass all three

fundamental aspects of a paradigm – ontology, epistemology and methodology – so it may not qualify as a paradigm shift in the strictest sense. However, it certainly challenges existing epistemological beliefs about how we understand and acquire knowledge of reality.

Traditional coding practices are based on the belief that researchers must first segment data, classify it into categories and subcodes, and derive themes. For some, these derived themes represent the final results, while others may probe further, querying the coded data for deeper insights. In contrast, proponents of the new vision argue against the necessity of these steps. They believe that engaging in a dialogue with AI, which can navigate through data more adeptly than humans, is sufficient for uncovering knowledge. This approach challenges conventional epistemological assumptions in qualitative research, suggesting that knowledge can be directly and dynamically accessed and understood through advanced AI interactions, rather than through the meticulous organization and analysis of data. This represents a significant shift in beliefs about how researchers can effectively comprehend and interpret reality.

Building on the argument about the potential shift in qualitative data analysis approaches, Thomas Kuhn's insights from 2012 resonate profoundly with the adoption of new technology. He stated, 'A decision between alternate ways of practising science is called for, and in the circumstances, that decision must be based less on past achievements than on future promise' (p. 153). This observation is particularly relevant as we consider integrating new, evolving technologies into research methodologies. The adoption of AI in qualitative analysis is a dynamic development, gradually gaining traction among researchers. We can further follow Kuhn's argument as he suggests that the initial endorsement of the new technology might often rest on subjective, even aesthetic considerations.

This new approach has already garnered its first wave of supporters who are crucial for its development and validation. Over time, as purpose-built tools are developed – beyond generic applications such as ChatGPT designed without the specific needs of qualitative analysts in mind – the effectiveness and utility of AI in qualitative research will become more apparent.

Paraphrasing Kuhn (2012, p. 157), even if the proponents of this new approach are few and their motives occasionally questioned, their competence in refining and exploring the approach will demonstrate its potential. As these innovators continue to experiment with and refine AI-driven methodologies, they will delineate what it means to be part of a community that embraces these advanced tools. Consequently, the arguments supporting this new approach will grow stronger and more persuasive, potentially leading to its broader acceptance and integration into standard qualitative research practices. This evolutionary process underscores the transformative power of innovation

in reshaping scientific inquiry, illustrating how a new technology can redefine the landscape of research methodologies.

7.5 Conclusion

As demonstrated throughout this chapter, the integration of AI not only advances technological capacities but also compels a re-evaluation of established analytical practices within qualitative research. This re-evaluation is critical, as it highlights the potential for AI to deepen and expand our understanding of qualitative data, necessitating a shift in the methodological approaches traditionally embraced by the field.

The historical trajectory of coding in qualitative analysis, as detailed in this chapter, shows a disciplined evolution towards more structured and systematic methods. However, the advent of AI technologies introduces anomalies that challenge these established methods, echoing the phases of scientific revolutions described by Thomas Kuhn. These anomalies do not signify faults per se but represent emerging challenges and opportunities that the current methodological standards in qualitative research must address.

The way AI technology is currently implemented on the one hand, and the resistance encountered from the research community towards integrating AI tools on the other hand reflects a broader hesitancy to alter deeply entrenched practices. Yet, as this chapter argues, such integration is less about replacement and more about enhancement. AI's potential to facilitate and enrich the analytical process illustrates the necessity for an adaptive approach that accommodates new tools while reconsidering methodological foundations.

In conclusion, the incorporation of AI into qualitative data analysis is not just about technological advancement but also about the willingness of the field to embrace a new method that could expand and deepen our understanding of qualitative data. This shift requires an openness to re-evaluating established practices and a bold vision for the future – a vision that the author of this chapter is committed to developing and demonstrating through innovative AI-based tools. This, in Kuhnian terms, is how a revolution begins.

References

Adduesselam, M.S. (2023) Qualitative data analysis in the age of artificial general intelligence. *Journal of Advanced Research in Qualitative Data Analysis* 1, 10–25.

Adorno, T.W., Frenkel-Brunswik, E., Levinson, D.J. and Sanford, R.N. (1950) *The Authoritarian Personality* . Harper C Brothers, New York.

Bazeley, P. (2013) *Qualitative Data Analysis: Practical Strategies* . Sage Publishing, London.

Becker, H.S. and Geer, B. (1960) Participant observation: the analysis of qualitative field data. In: Adams, R.N. and Jack, J.P. (eds) *Human Organization Research* . Dorsey Press, Homewood, Illinois, pp. 267–289.

Becker, H.S., Hughes, E.C., Geer, B. and Strauss, A.L. (1961) *Boys in White: Student Culture in Medical School* . University of Chicago Press, Toronto, Canada.

Becker, H.S., Gordon, A.C. and LeBailly, R.K. (1984) Fieldwork with the computer: criteria for assessing systems. *Qualitative Sociology* 7, 16–33.

Bird, A. (2022) Thomas Kuhn. In: Zalta, E.N. (ed.) *The Stanford Encyclopedia of Philosophy* , spring 2022 edition. Available at: https://plato.stanford.edu/archives/spr2022/entries/thomas-kuhn/ (accessed 24 February 2025).

Braun, V. and Clarke, V. (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 77–101.

Braun, V. and Clarke, V. (2023) Toward good practice in thematic analysis: avoiding common problems and be(com)ing a knowing researcher. *International Journal of Transgender Health* 24, 1–6.

Chapoulie, J.-M. (1987) Everett C. Hughes and the development of fieldwork in sociology. *Urban Life* 15, 259–298.

Charmaz, K. (2014) *Constructing Grounded Theory* , 2nd edn. Sage Publications, Los Angeles, California.

Christou, P.A. (2024) Thematic analysis through artificial intelligence (AI). *The Qualitative Report* 29, 560–576.

Corbin, J. and Strauss, A. (2015) *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* , 4th edn. Sage Publications, London.

Feyerabend, P. (1975) *Against Method: Outline of an Anarchistic Theory of Knowledge* . New Left Books, London.

Fielding, N.G. and Lee, R.M. (1998). *Computer Analysis and Qualitative Research* . Sage Publications, London.

Friese, S. (2019) *Qualitative Data Analysis with ATLAS.ti* , 3rd edn. Sage Publications, London.

Gamieldien, Y., Case, J.M. and Katz, A. (2023) *Advancing qualitative analysis: an exploration of the potential of generative AI and NLP in thematic coding*. SSRN. Available at: https://ssrn.com/abstract=4487768 (accessed 24 February 2025).

Gao, J., Tsu, K., Choo, W., Cao, J., Lee, R.K.W. *et al*. (2023) CoAIcoder: examining the effectiveness of AI-assisted human-to-human collaboration in qualitative analysis. *arXiv preprint* . Available at: https://doi.org/10.48550/arXiv.2304.05560 (accessed 26 September 2024).

Glaser, B.G. and Strauss, A.L. (1967) *The Discovery of Grounded Theory: Strategies for Qualitative Research* . Aldine Publishing Company, Chicago, Illinois.

Goyanes, M., Lopezosa, C.L. and Jordá, B. (2024) Thematic analysis of interview data with ChatGPT: designing and testing a reliable research protocol for qualitative research. SocArXiv preprint. Available at: (accessed 26 September 2024).

Hitch, D. (2023) Artificial intelligence (AI) augmented qualitative analysis: the way of the future? SSRN. Available at: https://ssrn.com/abstract=4451740 (accessed 26 September 2024).

Hyman, H.H. and Sheatsley, P.B. (1954) The authoritarian personality: a methodological critique. In: Christie, R. and Jahoda, M. (eds) *Studies in the Scope and Method of "The Authoritarian Personality": Continuities in Social Research* . Free Press, Glencoe, Illinois.

Jiang, J.A., Wade, K., Fiesler, C. and Brubakerm, J.R. (2021) Supporting serendipity: opportunities and challenges for human–AI collaboration in qualitative analysis. In: *Proceedings of the ACM on Human–Computer Interaction, Vol. 5. Association for Computing Machinery, New York, Article S4*.

Kálmán, A. (2016) The meaning and importance of Thomas Kuhn's concept of paradigm shift: how does it apply in education? *Opus et Educatio* 3, 96–107.

Kuckartz, U. (2014) *Qualitative Text Analysis: A Guide to Methods, Practice and Using Software* . Sage Publications, London.

Kuckartz, U. and Rädiger, S. (2024) Qualitative Inhaltsanalyse. Methoden, Praxis, Umsetzung mit Software und *künstlicher Intelligzen*. Beltz Juventa, Weihnheim (English version, chapter 10: Integrating Artificial Intelligence (AI) in Qualitative Content Analysis. Available at: https://doi.org/10.1177/16094069231 211248 (accessed 24 February 2025).

Kuhn, T.S. (1962) *The Structure of Scientific Revolutions* . University of Chicago Press, Chicago, Illinois.

Kuhn, T.S. (2012) *The Structure of Scientific Revolutions* , 50th edn. University of Chicago Press, Chicago, Illinois.

Lakatos, I. and Musgrave, A. (1970) *Criticism and the Growth of Knowledge* . Cambridge University Press, Cambridge, UK.

Mayring, P. (2022) *Qualitative Inhaltsanalyse: Grundlagen und Techniken* , 13th edn. Beltz, Weinheim, Germany.

Miles, M.B., Huberman, A.M. and Saldaña, J. (2019) *Qualitative Data Analysis: A Methods Sourcebook* , 4th edn. Sage Publications, London.

Morgan, D.L. (2025) Query-based analysis: a strategy for analyzing qualitative data using ChatGPT. *Qualitative Health Research*, forthcoming.

Morgan, D.L. (2023) Exploring the use of artificial intelligence for qualitative data analysis: the case of ChatGPT. *International Journal of Qualitative Methods* 22. Available at: (accessed 25 February 2025).

Nguyen-Trung, K. (2024). ChatGPT in thematic analysis: can AI become a research assistant in qualitative research? OSF.io preprint. Available at: https://osf.io/preprints/osf/vefwc (accessed 25 February 2025).

Palmer, V. (1928) *Field Studies in Sociology: A Student's Manual* . Chicago University Press, Chicago, Illinois.

QSF Mailing List (n.d.) Mailingliste Qualitative Sozialforschung. Archive. Available at: https://lists.fu-berlin.de/listinfo/qsf_l#archive (accessed 25 February 2025).

Richards, L. (2020) *Handling Qualitative Data: a Practical Guide* , 4th edn. Sage Publications, London.

Richards, L. and Richards, T. (1991) The transformation of qualitative method: computational paradigms and research processes. In: Fielding, G. and Lee, R.M. (eds) *Using Computers in Qualitative Research* . Sage Publications, London, pp. 38–53.

Saldaña, J. (2021) *The Coding Manual for Qualitative Researchers* , 4th edn. Sage Publications, London.

Sṃen, M., Sṃen, Sṃ.N. and Sṃahin, T.G. (2023) A new era for data analysis in qualitative research: ChatGPT! *Shanlax International Journal of Education* 11 (Suppl. 1), 1–15.

Silver, C. and Lewins, A. (2014) *Using Software in Qualitative Research: A Step-by-step Guide* , 2nd edn. Sage Publications, London.

Strauss, A. and Corbin, J. (1998) *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* , 2nd edn. Sage Publications, Thousand Oaks, California.

Theelen, H., Vreuls, J. and Rutten, J. (2024) Doing research with help from ChatGPT: promising examples for coding and inter-rater reliability. *International Journal of Technology in Education* 7, 1–18.

Whyte, W.F. (1955) *Street Corner Society* . University of Chicago Press, Chicago, Illinois.

Xiao, Z., Yuan, X., Liao, Q.V., Abdelghani, R. and Oudeyer, P.-Y. (2023). Supporting qualitative analysis with large language models: combining codebook with GPT-3 for deductive coding. In: *Proceedings of the 28th International Conference on Intelligent User Interfaces Companion (IUI '23 Companion), 27–31 March 2023, Sydney, New South Wales, Australia* . Association for Computing Machinery, New York, pp. 75–78.

Zhang, H., Wang, C., Xu, J., Li, Y., Ji, J. *et al*. (2023) Redefining qualitative analysis in the AI era: utilizing ChatGPT for efficient thematic analysis. *arXiv preprint*. Available at: https://arxiv.org/abs/2309.10771v2 (accessed 25 February 2025).