Qualitative Health Research
2024, Vol. 34(7) 595–606

DOI: 10.1177/10497323231217392
journals.sagepub.com/home/qhr

Artificial Intelligence Augmented Qualitative Analysis: The Way of the Future?

Danielle Hitch[1,2]

Abstract

The artificial intelligence (AI) revolution is here and gathering momentum, thanks to new models of natural language processing (NLP) and rapidly increasing adoption by the public. NLP technology uses statistical analysis of language structures to analyse and generate human language, using text or speech as its source material. It can also be applied to visual mediums like images and videos. A few qualitative research early adopters are beginning to adopt this technology into their work, but our understanding of its potential remains in its infancy. This article will define and describe NLP-based AI and discuss its benefits and limitations for reflexive thematic analysis in health research. While there are many platforms available, ChatGPT is the most well-known and accessible. A worked example using ChatGPT to augment reflexive thematic analysis is provided to illustrate potential application in practice. This article is intended to inspire further conversation around the role of AI in qualitative research and offer practical guidance for researchers seeking to adopt this technology.

Keywords

ChatGPT, artificial intelligence, reflexive thematic analysis, qualitative analysis

Introduction

Artificial intelligence (AI) has been around for many years and is an established presence in healthcare (Briganti & Le Moine, 2020). However, the accelerating developments in this field have bought wider recognition of both its potential positive and negative consequences. AI has moved from being a niche area of interest for the technologically savvy to a mainstream concern in a matter of months.

Some believe these developments are part of a fourth industrial revolution (Oosthuizen, 2022), which will forever change the way we work in all sectors of the economy – including academic research.

Despite the many potential advantages identified for AI, its adoption also introduces significant ethical dilemmas. All chats and other information uploaded to AI platforms are used to develop the underlying algorithms that enable computer learning. In some cases (such as ChatGPT), inputted information is also in the public domain, leading to some academics to describe it as a 'data privacy nightmare' (Gal, 2023). There is also an ongoing debate in the literature around whether ChatGPT can be considered a co-author or not for scholarly publications (Rahman et al., 2023). Amid the enthusiasm for what AI can offer humanity, there are also many calls to proceed with caution.

To date, the use of AI in qualitative research has been somewhat limited, with most researchers relying on human-driven qualitative analysis methods. However, some researchers have recognised the potential role AI could play in making traditionally time-consuming analysis tasks more efficient and less burdensome. Cheligeer et al. (2022) present a case study example of applying natural language processing (NLP) to open-ended survey responses from people experiencing chronic pain. They compared a traditional approach to coding as part of the initial stages of thematic analysis with two forms of NLP for the analysis of interviews

[1]Deakin University, Geelong, VIC, Australia
[2]Western Health, Sunshine Hospital, St Albans, VIC, Australia

Corresponding Author:
Danielle Hitch, Deakin University, Waterfront Campus, Geelong, VIC 3217, Australia.
Email: dani.hitch@deakin.edu.au

conducted as part of a larger public health study. After transforming their data into a computer-readable format, they used open-source software to automate the process of identifying themes and exploring the relationships between them. They found the technology enabled a more efficient and accurate approach to analysis but highlighted there are also limitations to the technology which should be considered.

Prior to the widespread availability of AI platforms, Lennon et al. (2021) developed and tested an automated qualitative assistant (AQUA) software tool based on NLP to automate coding and analysis of qualitative data. The data was derived from open free text responses to a survey and use the software to undertake a rapid descriptive thematic analysis, an approach that presents a thematic clustering of the findings rather than thematic schemas or interpretations (Kamalakannan et al., 2021). This software proved to be highly accurate in its identification of themes and subthemes, and the authors were able to quantify a 75% reduction in the time required for coding of data. In addition, they found the AQUA tool had provided insights into relationships between key concepts that they themselves had not identified. Similar to Cheligeer et al. (2022), they concluded the benefits of AI for qualitative analysis outweigh its limitations.

There is only one description of the application of AI within a qualitative research methodology. A study of the lived experience of men experiencing infertility (Osadchiy et al., 2020) utilised an NLP approach supported by the data analytic tool BigQuery and the textual analysis program Language Inquiry and Word Count. Descriptive thematic analysis was undertaken on 133 Reddit posts, which were subjected to both thematic and semantic analysis. In this case, AI was employed to work through the chosen methodologies, rather than as a methodology in itself.

This article will provide further detail and discussion around the use of machine learning algorithms and NLP in reflexive thematic analysis. A definition and description of NLP-based AI is provided as foundational knowledge. All the existing studies applying AI based on NLP to qualitative analysis have conducted thematic analysis; however, their approaches appear more like descriptive thematic analysis than the more reflective forms of this analytical approach (Braun & Clarke, 2021a). Therefore, the application of AI to reflexive thematic analysis is yet to be explored in the literature. A worked example using ChatGPT within a process of reflexive thematic analysis is provided to illustrate just one of many potential applications of AI. Finally, the benefits and limitations of AI for reflexive thematic analysis are discussed. I will argue that AI can augment, but not replace, the work of human researchers in producing rigorous reflexive thematic analysis, so long as it is applied in a mindful and ethical manner. This discussion is intended to inspire further conversation around the role of AI in qualitative research and offer practical guidance for anyone seeking to adopt this technology in their work.

What Is Artificial Intelligence?

While AI is rapidly becoming part of everyday life, it is a new concept to many people and general knowledge about the field is still developing. A German study of approximately 2000 adults (Cornelius et al., 2022) found less than a quarter of participants could accurately describe the key features of AI, and the majority of them held a neutral attitude towards its use in healthcare. However, many also expressed fears about the potential of AI to replace human workers in the longer term and the potential for breaches of privacy. An Australian study (Yigitcanlar et al., 2022) of 605 city dwellers also highlighted that trust in AI depends on who is deploying the technology, with commercial entities and government agencies considered the least trustworthy. I therefore make no assumptions about the readers understanding of AI technology and will begin by introducing some of the key concepts in this field.

Artificial intelligence involves the "development of computer systems capable of performing tasks that typically require the features of human intelligence, such as visual perception, speech recognition, decision-making, and problem solving" (Russel & Norvig, 2016). Algorithms are mathematical rules and procedures that instruct computers to execute tasks and solve problems based on input provided by human users (Cormen et al., 2009). The key feature of AI is that these algorithms also allow the computers to learn from the inputs provided and use this learning to make better decisions in the future (Russel & Norvig, 2016). New algorithms are being developed constantly as computers master increasingly complex problems, and this is driving the rapid acceleration in this technology.

There are two categories of AI which reflect their differing abilities to manage complex tasks. Narrow AI (also known as weak AI) performs a restricted range of simple and straightforward tasks and relies primarily on algorithms that use pattern recognition for prediction (Ng & Leung, 2020). An example of narrow AI using machine learning is speech recognition, as deployed by home assistants such as Siri, Alexa, and Google Home. The algorithms which drive general AI (also known as strong AI) enable the computer to autonomously problem solve and therefore learn from the inputs it receives (Stahl, 2021). General AI aims to enable computers to mimic human interactions and reasoning (Kerzel, 2020), potentially leading to computers achieving equal capabilities to humans.

Superintelligence is a hypothetical (but emerging) type of AI that seeks to surpass human intelligence (Baum, 2018). The ability of AI to autonomously learn and become 'more like us' over time is provoking increasing ethical debate and calls for a pause in development considering its consequences for the world (Future of Life Insitute, 2023).

The forms of AI currently available with the best potential to augment qualitative research are general AI platforms which utilise NLP. NLP is an area of AI derived from linguistics and based on statistical analysis of language structure which focuses on interactions between computers and human languages (Kanaparthi, 2022). It involves machine learning, learning models, and algorithms to allow computers to analyse the complete meaning of text or speech, including intent and emotions (Chowdhury & Nath, 2001). Computers can then interpret these inputs and generate human-like language in response; however, the contextuality and ambiguity of human languages remains a recognised and ongoing challenge in this area of AI. NLP algorithms cannot always account for or understand these nuances, particularly when the algorithm has little data available on the language or dialect in question.

Effective Use of NLP-Based AI

There is a plethora of AI platforms (and associated extensions, plug-ins, or add-ons) available, and more are appearing every day. The tool used in this discussion is Chat Generative Pre-Trained Transformer (GPT) which is an open access chatbot which produces human-like conversations. It was launched on November 30, 2022, and is currently visited by 25 million people each day (Nerdynav, 2023). The tool uses a combination of machine learning, deep learning, and NLP techniques to execute a variety of tasks, including translation, question answering, summarisation, plan generation, and conversational interaction.

ChatGPT has been trained on an immense and diverse trove of text data including books, articles, websites, and social media posts. When you ask ChatGPT a question or ask it to generate an output, it does not conduct an Internet search and present the answer like a search engine. Instead, the platform breaks down human inputs into smaller chunks (like words and phrases) and then uses models and algorithms to identify information it deems as relevant from what the platform has learnt to date. It then uses algorithms to generate a coherent response which it judges to be an appropriate answer in natural language that mimics human communication. The initial phase of training lays a foundation with the computer completing unsupervised learning from a large corpus of text. Its learning is then fine-tuned via supervised reinforcement and adjustment with more specific selections of text (Greengard, 2022).

Text is inputted by users for analysis by the algorithms which identify both the most relevant information and its overall context. A key aspect of ChatGPT (and other AI tools) is the quality of the data inputted – what you get out depends on what you put in. These inputs are called prompts, and they are used in general conversation to provide the computer contextual information. Prompts are entered as text, and the user engages with the chatbot in a conversational manner. Prompts are not limited to the posing of questions; they can include information about the format of the output you are seeking (i.e. style of writing and intended audience), dot points about key topics you want addressed, instructions on the perspective taken in the response, and specific requirements (i.e. no jargon and word count) (Scharth, 2023). The prompts also contribute to the computer learning process and drive improvements in the accuracy and quality of responses. Millions of prompts have contributed to the updated ChatGPT4 released on March 15, 2023, which is far more capable across all tasks than the original version (Popli, 2023). For example, ChatGPT can now identify academic citations relevant to the output with increasing accuracy, whereas it previously generated apparently random citations.

The approach to selecting prompts, therefore, has much in common with a process qualitative researchers are already familiar with – choosing Boolean operators when searching electronic databases. Precise terms improve the accuracy of outputs and return the relevant information more efficiently. The combination of terms used enables the construction of complex prompts, providing the computer with important contextual or modifying information. Guides to writing 'good' prompts are readily available online, although the conversational nature of ChatGPT makes it an accessible interface that requires no special training or skills. However, the emergence of the new and very highly paid job title of 'prompt engineer' (Al-Sibai, 2023) indicates the value of upskilling in this area if you want to use AI regularly.

Worked Example of AI Augmented Reflexive Thematic Analysis

To illustrate how AI (in this case ChatGPT) can augment qualitative analysis, I selected a feature newspaper article about long COVID in Australia published in *The Age* newspaper of Melbourne (Souter, 2023). The feature included interviews with both patients and health experts and focused on the national response to this emerging health problem. AI had been applied to qualitative data collected from open answer survey responses and social media posts in previous research (Cheligeer et al., 2022; Osadchiy et al., 2020), but I wanted to explore its applications for semi-structured interviews and other long-form source material. The research questions I sought to answer were (1) How do Australian health experts and people with long COVID describe healthcare system responses to long COVID? and (2) What are the similarities and differences between the perspectives of Australian health experts and people with long COVID?

To analyse this data, I chose to employ reflexive thematic analysis. Reflexive thematic analysis was developed by Virginia Braun and Victoria Clarke as a flexible approach to code development, and it can be applied within various qualitative approaches (Braun & Clarke, 2021a). Reflexive thematic analysis is a qualitative research method which analyses and interprets patterns in textual data and emphasises the active role that researchers play in interpreting themes within their unique context (Braun & Clarke, 2021a). Reflexivity refers to routine reflection on researchers' assumptions, expectations, choices, actions, values and politics, and their potential influence on every step of the analytical process (Braun & Clarke, 2021b). Researchers are required to iteratively reflect throughout the analytical process as new understanding and interpretations are developed through inductive engagement with participants and their data. This worked example was informed by the six stages of this analytical process, and the following example illustrates the role of ChatGPT throughout. I am far from an expert user of the platform, and therefore, this example reflects a novice approach to its use.

Familiarisation

At this stage, the researcher must immerse themselves in their data, reading and re-reading their source material to gain a holistic understanding of the content and its context (Braun & Clarke, 2021b). ChatGPT can be used to generate summaries of source material, which can support the rapid identification of key content and sentiments. These prompts can also generate summaries to a specified level of detail, for example, a word count (see Supplementary Material 1 for a high-level summary of the article contents). While you could enter more specific prompts to generate a more detailed summary, this approach doesn't allow the deep immersion required to fully engage with source material. AI would therefore play a minor role at this stage of reflexive thematic analysis, as reliance on it would prevent researchers from gaining the deep understanding of the richness in their data which underpins the rest of this analytical approach.

Coding

It is at the coding stage that AI really begins to come into its own. Coding involves the researcher highlighting meaningful segments of text and may be based on either an open inductive process or pre-determined coding frameworks depending on the approach to thematic analysis adopted (Braun & Clarke, 2021c). At this point, it's important to reflect on the quality of the input you provide to ChatGPT and consider whether your data may require 'cleaning' or 'transformation' to make it easier for the platform to understand. The second research question required a comparative analysis between two participant groups, and so the source material from the article was reformatted into separate transcripts for Australian health experts and people with long COVID. Extraneous material related to the journalist's reporting (i.e. descriptions of location, the journalist's response, or their interpretation) was also removed, to provide the computer with a more target and focus set of data. A new chat was initiated for the analysis of each transcript, to avoid ChatGPT becoming confused about which data was being referred to in each prompt.

To enable automated coding, the source material must first be uploaded to ChatGPT. There are character limits for each input and response, although this is likely to disappear in future as ChatGPT continues to evolve. If the limit is exceeded, the input or response cuts off mid-sentence. The ChatGPT Splitter extension (and others) can parse text into chunks for easier upload. After each chunk, ChatGPT will automatically start to generate a response, but this can be halted by concluding the text with instructions (i.e. *Reply with OK only*). This process continues until the entire source material has been inputted.

Once uploaded, coding can begin. In this case, multiple people with long COVID contributed to the article (not all of whom were individually identified), and therefore, coding was occurring at the group level, like a focus group transcript. However, the application of AI to interview transcripts would enable detailed comparison of key codes between individual participants. I referred to themes in the ChatGPT prompt for this phase of analysis, as the term 'codes' seemed confusing for the platform. The codes identified are presented in Table 1, along with my independently completed coding of the source data.

Developing Initial Themes

Once codes are generated, the researcher begins to identify patterns and relationships within the data and develop preliminary themes (Braun & Clarke, 2021b). As such, themes are developed from the codes. Independent coding by two researchers is considered best practice to

Table 1. What Are the Key Concepts Within the Data Inputted Above?

ChatGPT prompt	Author codes
People with long COVID	
1. Long COVID	1. Symptoms (brain fog, breathlessness, fatigue, muscle aches and heaviness, chronic gastrointestinal problems, postural orthostatic, tachycardia syndrome or POTS, allergies, and difficulties speaking)
2. Cognitive impairment	2. Feelings of desperation, grief, and frustration
3. Brain fog	3. Lack of knowledge and support from healthcare providers and family
4. Physical symptoms (fatigue, muscle aches, chronic gastrointestinal problems, POTS, and allergies)	4. Impact on activities of daily life (i.e. employment, finances, meal preparation, reading, exercising, mobility, and housework)
5. Emotional impact of desperation, grief, and loss of identity	5. The future – uncertainty, changed expectations, and new ways of doing things
6. Lack of support and understanding	
7. Financial burden	
8. Uncertainty about recovery	
9. Change and adaptation	
10. Risk of long COVID	
Health experts	
1. Long COVID	1. Symptoms (brain fog, breathlessness, fatigue, muscle weakness, exercise intolerance, breathlessness, and mental health issues)
2. Symptoms	2. Still investigating potential mechanisms (brain changes, viral persistence, immune responses, microclots, and interferons)
3. Brain fog	3. Vaccine as a mitigator or preventor
4. Fatigue	4. Perceptions of how patients deal with impact on daily life
5. Cognitive difficulties	5. Uncertainty about the best treatment or therapy
6. Abnormal MRIs	6. Limited service access and funding for hospital-based (long COVID clinics) and community-based (GP) services
7. Impact on daily activities	7. It's a real condition and patients should be believed
8. Vaccination	
9. Uncertainty	
10. Coping skills	
11. Immune response	
12. Interferons	
13. Therapeutic interventions	
14. Clinical trials	
15. Postcode-driven services	
16. Funding	

support trustworthiness in some approaches to qualitative data (Sweeney et al., 2013) but is not necessarily utilised in reflexive thematic analysis. In this case, myself and ChatGPT could be interpreted as multiple coders and therefore an effort to increase the 'reliability' of the analysis (Braun & Clarke, 2021a). However, my intention was not to reduce bias but rather to compare the two analyses as part of the reflexive process. I began by comparing the codes generated by ChatGPT with my own, reflecting on the similarities and differences. ChatGPT identified long COVID as an overall concept, while I assumed this was the overall subject and therefore did not identify it as a code. However, generally speaking, the codes identified were very similar. However, subtle but identifiable differences were also evident. Both ChatGPT and myself identified symptoms as a key code in both participant groups; however, I also identified some of the less commonly reported symptoms of long COVID in the transcript from people with long COVID. ChatGPT also separated symptoms into physical and cognitive categories, while I took a holistic perspective which was more broadly inclusive of symptom concepts. Identifying the emotional impact of long COVID as a separate code seemed appropriate, as while it could be interpreted as just another symptom, this concept was more strongly present within the source material. Finance was the only activity of daily living identified by ChatGPT, and I believe my professional background as an occupational therapist influenced my identification of more data related to this code. This is perhaps indicative of the overall difference in coding focus between ChatGPT (who identified more descriptive or concrete codes) and the author (who tended to subsume several into broader concepts). This observation suggests AI developed themes are more 'literal' in comparison to the reflexive thematic approach I was taking by engaging with their latent or implicit meaning (Braun & Clarke, 2021a).

In two cases, codes were identified by one coder but not the other. Risk of long COVID was a concept only identified by ChatGPT while 'It's a real condition and patients should be believed' was only identified by myself. As would occur when two human researchers collaborate, these discrepancies prompted me to conduct further reading and notation of the transcripts. The transcript was also interrogated with additional prompts posed separately to each transcript (see Supplementary Material 2).

While the first section and last sentence of the ChatGPT output for 'risk of COVID' were generally descriptive of long COVID, content in the mid-section of the output closely mirrored my notes. Following this review, I agreed this code was evident in the data and added a code called 'Long COVID could happen to anyone (hospitalised or not, rampant in the community, reinfection)'. Upon returning to the health expert transcript, I saw there were data suggesting this participant group also had difficulty being believed (at least in the early days of the pandemic). This led to a revision of this code to 'growing awareness that long COVID is real'. In summary, ChatGPT identified a greater number of codes; however, this process of comparison and reflection demonstrated they were mostly aligned with the broader codes that I identified. Both sets of codes therefore contributed to the development of preliminary themes (see Supplementary Materials 3).

Reviewing Themes

The next phase of reflexive thematic analysis involves reviewing and refining preliminary themes to ensure they are grounded adequately in the data and expressed coherently (Braun & Clarke, 2021c). I revisited the source material to check the fit between the preliminary themes and the data and considered if any could be merged before drafting revised definitions for each theme. While none of the themes identified for people with long COVID seemed suitable for merging, themes from the health expert data appeared to offer an opportunity for consolidation. I identified the theme of 'Long COVID service delivery' as potentially uniting both 'uncertainty about best care' and 'barriers to services and funding'. As shown in Table 2, I tested the relationship between these two codes in ChatGPT whose response supported the decision to merge them together.

My initial assessment was the preliminary themes were clearly evident within the source material, and I sought to confirm this using ChatGPT. The coherence of each theme was tested with the following prompt: '*Summarise the data in the input above relevant to the following theme [THEME]*' (see Supplementary Material 4 for all summaries and the finalised themes). The summaries of each theme generated by ChatGPT and the author were very similar.

Refining Themes

The penultimate step in reflexive thematic analysis is to finalise the definitions and names of each theme and identify quotes which illustrate their meaning to participants (Braun & Clarke, 2021c). While the themes distinctive to each group had already been defined and named in the previous step, those common to both transcripts were now developed using the same process (see Supplementary Materials 5 for the ChatGPT theme summary, my corresponding theme summary, and the finalised theme name). To find illustrative source material, I initially identified two to three quotes for each theme from my readings of the transcripts. ChatGPT was then prompted to '*List all quotes from [data source] which illustrate the theme of [theme name]*'. All of the quotes I

Table 2. Testing Decision-Making for Potential Code Merger.

ChatGPT prompt	ChatGPT output
Does the data about the themes of 'uncertainty about best care' and 'lack of service access and funding' in this input address the concept of service delivery?	Yes, the data about the themes of 'uncertainty about best care' and 'lack of service access and funding' in the post above addresses the concept of service delivery. Service delivery refers to how healthcare services are provided and accessed by patients, including factors such as quality of care, access to resources and funding, and the availability of dedicated clinics or facilities. The themes of uncertainty about best care and lack of service access and funding can have a significant impact on service delivery for long COVID patients and can contribute to the challenges and limitations they face in accessing appropriate care and support.

identified appeared in the lists generated by ChatGPT, along with others which tended to be associated with the theme but not as directly relevant.

The final task in this phase of analysis involved a comparison between the two participant groups to answer the second research question. I produced a preliminary list of themes common to data from both people with long COVID and health experts, complemented by a list of themes distinctive to each group. The theme lists from each transcript were then uploaded to ChatGPT, with the prompt '*Above are themes which have emerged from interviews with people living with long COVID (list 1) and health experts (list 2). Please compare and contrast the two lists and discuss which of the themes are similar and which are different between the lists*'. The output of this prompt identified the same distinctive and shared themes as my original lists, and the analysis concluded with no need to return to the data for confirmation (Table 3).

Analytic Report

Writing an analytic report describing themes and findings is the final stage in the reflexive thematic analysis process (Braun & Clarke, 2021c); however, this was not the focus of this article. AI is increasingly used in the production of academic journal articles and research reports, as demonstrated by MacDonald et al. (2023) who used the platform to draft an academic paper.

However, they emphasise AI cannot fully automate the process of science dissemination, and critical analysis must be applied to all outputs generated.

Reflection on AI Augmented Reflexive Thematic Analysis

My experience of applying AI to reflexive thematic analysis identified several advantages and disadvantages for this approach. AI may identify patterns and themes in the data not immediately apparent to human researchers (such as the theme 'risk of long COVID'), leading to potentially novel insights. In the worked example, this prompted me to reflect on my personal stance and assumptions, which in turn prompted a return to the source material and further investigation. Reflexive thematic analysis is not a linear process (Braun & Clarke, 2021c), and my choice to return to an earlier stage of the process provided an additional opportunity to engage and develop insight. In this sense, ChatGPT supported iterative analysis by encouraging a deliberative approach and the development of shared meaning. Researchers adopting a reflexive thematic analytical approach often develop and evolve their analysis with participants or others – in this case, the 'other' was ChatGPT.

This potential benefit of AI augmented reflexive thematic analysis should not be mistaken as a means to increase 'objectivity'. AI has been proposed as a way of validating qualitative analysis, which reflects the detached perspectives of the neo-positivistic paradigm (Cresswell, 2014). Positivist approaches may therefore adopt AI more readily due to this perceived alignment to their underlying values. However, in most qualitative research methodologies "there is no dichotomy between subjectivity and objectivity, there is only a dynamic in-between" (van Wijngaarden et al., 2017, p. 1741). In the worked example, there was a dynamic interaction between my human subjectivity and the apparent objectivity of ChatGPT, which suggests AI offers the potential for more than a simple validation of codes and themes.

Without this potential to support deeper interpretative analysis, AI would have no role to play in reflexive thematic analysis. Active and iterative engagement by researchers with theoretical and philosophical contexts, analytical decision-making, and the coding process are the hallmarks of this method (Braun & Clarke, 2019). Braun and Clarke describe the themes resulting from these analyses as "analytic outputs developed through and from the creative labour of our coding" (2019, p. 594). AI is designed to learn from experience and is trained from the input of multiple human (and therefore subjective) users. However, ChatGPT is not capable of independent creativity as it does not possess consciousness. The concerns currently being raised about AI are largely founded on anxiety about computers becoming 'more like us' over time, even to the point of achieving sentience (Gilbert & Martin, 2022). The

Table 3. Mapping of Themes by Transcript Source.

People with long COVID only	Both groups	Health experts
The emotional impact of long COVID	Experiences and perceptions of symptoms	Potential causes of long COVID
An uncertain future	Negative impact on daily life	Preventing long COVID through vaccination
Long COVID could happen to anyone	Long COVID is real but there's little knowledge or support	Long COVID service delivery

development of AI augmented Clinical Reasoning Support Systems (CRSS) are already proposed in medicine, “combining human and artificial intelligence into hybrid intelligence, were both perform clearly delineated and complementary empirical tasks” (van Baalen et al., 2021, p. 526). So, while ChatGPT is not currently capable of independent reflexive thematic analysis, the boundary between human and non-human interpretations of data is likely to become blurrier over time.

So, what does AI offer researchers using reflexive thematic analysis? Chelingeer et al. (2022) reported AI augmentation enabled increased accuracy and efficiency for their descriptive qualitative analysis.

Efficiency can be a contentious concept in reflexive thematic analysis, given the tension between getting an analysis done and getting it done properly. Braun and Clarke (2021b) hold hope for the future of ‘slow’ thematic analysis, and I share their belief that high-quality analysis takes time and multiple returns to the data. However, they acknowledge that time can be a luxury and a privilege. Long COVID provides a great example of an urgent global health problem that really can’t wait until qualitative researchers get their act together. Insisting on detailed (and therefore time intensive) reflexive thematic analysis also suggests limited engagement in the context of the study, which is ironically at the very heart of qualitative analysis.

The neo-positivist definition of accuracy (as a source of reliability) is irrelevant to reflexive thematic analysis. However, researchers must still adopt a systematic approach to ensure their identification, development, and interpretation of patterns in the data engage critically with their own stance and presuppositions. AI may therefore reduce the risk of misinterpretation by providing an alternative analysis against which researchers can test, interrogate, and critique their own analysis. AI won’t make reflexive thematic analysis more ‘correct’, but it might support a deeper and more critical approach to reflexivity.

Another potential advantage of AI platforms like ChatGPT is their ability to work efficiently through large datasets. While the worked example was based on a single newspaper article, the process described above could equally be applied to multiple transcripts or source data. The main limitation to analysing large datasets is the time required to enter it onto the platform. While the speed of obtaining the resulting output may be slower for large datasets, it would still be infinitely quicker than any human researcher could achieve. For example, identifying initial codes and themes from 100 interview transcripts takes human researchers many days of reading, reflection, and analysis. ChatGPT could complete this ‘first pass’ of the data in a matter of minutes, as part of familiarisation, coding, or both of these stages in reflexive thematic analysis.

AI therefore may also enable efficiency by maximising resources and ensuring the creativity and skills of human researchers are put to their best use. Spending less time on coding reduces the person hours needed to complete studies, potentially increasing the responsivity of reflexive thematic analysis to urgent health issues by contributing initial findings to often rapidly evolving policy and planning decisions. AI augmentation could also free up researchers’ time for deepening insights and developing more nuanced interpretation when returning to the data as part of a slow reflexive thematic analysis. As such, AI is a potential tool for both ‘quick and dirty’ initial interpretation in fast moving areas and subsequent rich and deep analyses.

There is a role for both approaches to reflexive thematic analysis in the field of long COVID, where the priority must be to support our participants and their healthcare providers navigate this profoundly debilitating and disruptive syndrome. A relatively rapid analysis can incorporate researcher reflection on their assumptions, expectations, choices, actions, values and politics, and engagement with overall patterns in the data. Preliminary studies of broad themes have value both in terms of their potential for timely influence and as markers upon which subsequent deeper analysis can build. To paraphrase well-known graduate students’ advice, a good reflexive thematic analysis can be completed with AI augmented analysis; a great reflexive thematic analysis can be completed using slow reflexive thematic analysis; and a perfect reflexive thematic analysis is neither.

Decisions around the use of AI in reflexive thematic analysis are up to the individual researcher, their theoretical orientation, their philosophical beliefs, and their access to suitable technology. Braun and Clarke (2021a) have always asserted that reflexive thematic analysis is intended to be a flexible rather than a prescriptive approach to qualitative analysis, and researchers should also consider their overall methodology as part of these decisions. As a pragmatist with a strong commitment to implementation, I would encourage anyone considering the use of AI in reflexive thematic analysis to also contemplate who we serve as qualitative researchers. From my perspective, it’s not about us – it’s about the impact our research can have for our participants. If AI supports the use of a systematic approach and the quicker completion of preliminary analyses, then this enables us to achieve this goal.

If researchers choose to use AI to augment their reflexive thematic analysis, its application in practice will be dependent on researchers learning to use AI platforms in effective and efficient ways. Guides to using the platform in research are beginning to appear in the literature, such as a demonstration of its use to reformulate and respond to reporting recommendations while simulating an epidemiological study (Sanmarchi et al., 2023). Methodologically tailored guidelines for AI augmented analysis would support its wider uptake, while also promoting rigour and quality.

However, it is equally possible that the introduction of AI could herald a reduction in the qualitative research workforce. Why pay for a university-qualified human researcher when their place could be taken by a freely available AI platform? Many sectors of the job market are in the process of being transformed by AI, particularly those founded on tasks which can be automated. Academia is not alone in facing this transformation with a recent analysis by Goldman Sachs (Hatzius et al., 2023) estimating 300 million jobs and 18% of global work could be impacted by AI automation. If elements of reflexive thematic analysis (or any other qualitative methodology) are found to be amenable, qualitative researchers will need to 'skill shift' by developing new capabilities to stay relevant in the job market (Brynjolfsson & McAfee, 2014). In this context, becoming an expert user of AI moves from a niche or 'value add' skill to become a fundamental requirement for employment.

The role of AI augmented reflexive thematic analysis in enabling equity also has potential advantages and disadvantages. Freely available platforms like ChatGPT may enable the democratisation of knowledge and expertise, in a similar way to Google and other search engines (Halavais, 2013). Many of the electronic supports for reflexive thematic analysis (such as NVivo and Dedoose) require financial resources which are not readily available to all researchers. ChatGPT is freely accessible to anyone with an Internet connection, but this does not guarantee its ability to bridge the 'digital divide' caused by inequitable access to information technology (International Telecommunication Union, 2021). AI could be employed to support both progress and equity, and striking a balance between these imperatives will require deep engagement with the ethical dimensions of this technology.

The ethical ramifications of AI augmented reflexive thematic analysis are critical considerations. While ChatGPT may demonstrate less biases than previous language learning models, an analysis by Zhou et al. (2023) identifies ongoing ethical issues including poor understanding of multiple languages, generation of disinformation, and the limits posed by the inputs used for training. Some of these problems were also identified by Cheligeer et al. (2022) who noted biases and errors in AI algorithms are an important limitation to their use in reflexive thematic analysis. From a research perspective, ChatGPT is a public platform and the data being entered for analysis is also being used by the AI system to learn and improve its own operations.

All data uploaded to ChatGPT not already in the public domain must be thoroughly de-identified to preserve the privacy and confidentiality of participants. The level of de-identification required by most ethics committees is the simple removal of identifying details (such as names, places, or singular experiences). The tension between keeping a participant's data within the context and ensuring that the context does not breach their confidentiality is not a new issue and has previously been explored in relation to case study research (McDonnell et al., 2000). However, the fact that AI platforms are open to anyone in the world may provoke additional concerns about the risk of inadvertent identification. As the technology is more widely adopted, its compliance with national and international data and privacy regulations needs urgent clarification. For example, the use of ChatGPT here has implications for many provisions in the General Data Protection Regulation (GDPR) (European Parliament, 2016) applicable across Europe.

The publication of de-identified and aggregate research data is allowed within ethical approval processes; however, guidelines specific to AI are yet to appear. In my opinion, the use of AI platforms such as ChatGPT within the analytical process should be included in plain language statements, to ensure participants are fully informed of how and where their data will be employed. Experienced qualitative researchers should lead the development of standards for the ethically sound use of AI, to ensure best practices are put into practice as soon as possible.

A final consideration for AI augmented reflexive thematic analysis is the speed of uptake for this technology. I consider myself an archetypal 'early adopter', who eagerly embraces new technology ahead of the general population and is comfortable with taking risks with new approaches (Rogers, 1962). Early adopters of ChatGPT have been found to have strongly positive sentiments about its disruptive nature and exercising creativity (Haque et al., 2022). However, only a minority of the population are early adopters, with up to half of people belonging to the late majority and laggard categories (Rogers, 1962). Despite the rapid development of AI in recent months, there will be a long period of transition between human-based approaches and AI augmented approaches to reflexive thematic analysis. There may also be some researchers who do not adopt AI at all, due to philosophical orientation, preference, or a lack of access to technology. AI augmented reflexive thematic analysis should therefore be thought of as an adjunct to existing approaches and not a replacement.

Conclusion and Future Directions

From my experience to date, AI augmented reflexive thematic analysis holds more advantages than disadvantages provided it is utilised in a mindful and critically aware manner. Our relationship with AI is evolving at an incredibly rapid pace, as I have observed in my interactions with ChatGPT over time. Much to my surprise, I found myself referring to ChatGPT as 'they/them' in conversations with human colleagues, began adding 'please' to my inputs, and automatically replied 'thank you' to a few of the outputs. However, a previous study found that people tend to be less open, agreeable, extroverted, and conscientious during interactions with AI (Mou & Xu, 2017).

I now think of ChatGPT as a virtual colleague, a novice in reflexive thematic analysis who nevertheless can make a useful contribution to less sophisticated aspects of analysis while they are learning and developing their skills. It became clear from the worked example that ChatGPT is not currently capable of the contextual interpretation and reflective deliberations required by reflexive thematic analysis. Human researchers remain the only ones who can interpret the meaning behind complex data and identify themes that reflect the holistic context of participants and phenomena. The algorithms of currently available AI simply cannot understand how social dynamics, cultural meanings, and unique circumstances impact the interpretation of reflexive thematic analysis.

ChatGPT itself acknowledges this, as demonstrated by the following output (dated 08/05/2023) when asked about the role of AI platforms in qualitative analysis:

> As an AI language model, I cannot perform qualitative analysis on the data above as it requires a more nuanced understanding of context, and the ability to interpret and analyze human language in context. Qualitative analysis involves identifying themes, patterns, and meanings in data through a systematic and subjective approach. It typically requires a human researcher with expertise in qualitative research methods and the specific field of study.

Clearly, the rapid evolution of AI has significant implications for current and future generations of all qualitative researchers and not just those using a reflexive thematic analysis approach. It would be unrealistic to expect novice reflexive thematic analysts to abstain from using AI until they 'learn their craft', and therefore, senior researchers must upskill in its use to ensure their students and mentees are properly trained in its effective and ethical application. Qualitative research is often challenging, and novice researchers could potentially see AI as a 'short cut' that reduces the effort required to engage with reflexive thematic analysis. However, as clearly demonstrated in the worked example, AI can augment but not adequately replace human researchers.

As such, the role of critical thinking and analysis is now even more crucial to qualitative analysis. Researchers much critically engage with both the advantages and disadvantages of platforms like ChatGPT to understand what they 'can' and 'cannot' do with reflexive thematic analysis. Researchers must also maintain a critical stance towards outputs and assess each on their relative merits. Within the worked example, I chose to incorporate some ChatGPT outputs into the analysis but discarded others that were not assessed to be relevant to the context of the data. Approaching data from a critical perspective is nothing new to qualitative researchers – we do it all the time with transcripts, conversations, and other sources.

From my perspective, the question of AI augmented reflexive thematic analysis is not whether it should be adopted but how to best adopt it. The revolution is here, and it has the potential to significantly enhance and transform the work of qualitative researchers from multiple traditions and approaches. AI can play an augmentative or supportive role in reflexive thematic analysis, but critical analysis remains the fundamental skill required to do high-quality research. The day all qualitative analysis will be fully automated by AI platforms remains in the far distant future for now … but that is not to say it will never arrive.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Danielle Hitch https://orcid.org/0000-0003-2798-2246

Supplemental Material

Supplemental material for this article is available online.

References

Al-Sibai, N. (2023). *Companies are paying bonkers salaries for people good at ChatGPT - "Expert prompt engineers can name their price"*. The Byte. https://futurism.com/the-byte/companies-paying-high-salaries-chatgpt

Baum, S. D. (2018). Countering superintelligence misinformation. *Information*, *9*(10), 244. https://doi.org/10.3390/info9100244

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, *11*(4), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

Braun, V., & Clarke, V. (2021a). Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and Psychotherapy Research*, *21*(1), 37–47. https://doi.org/10.1002/capr.12360

Braun, V., & Clarke, V. (2021b). The ebbs and flows of qualitative research: Time, change and the slow wheel of interpretation. In B. C. Clift, J. Gore, S. Gustafsson, S. Bekker, I. C. Batlle, & J. Hatchard (Eds.), *Temporality in qualitative inquiry: Theories, methods, and practices*. Routledge.

Braun, V., & Clarke, V. (2021c). *Thematic analysis: A practice guide*. Sage.

Briganti, G., & Le Moine, O. (2020). Artificial intelligence in medicine: Today and tomorrow. *Frontiers of Medicine*, *7*, 27. https://doi.org/10.3389/fmed.2020.00027

Brynjolfsson, E., & McAfee, A. (2014). *The second machine Age: Work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton and Company.

Cheligeer, C., Yang, L., Nandi, T., Doktorchik, C., Quan, H., Zeng, Y., & Singh, S. (2022). Natural language processing (NLP) aided qualitative method in health research. *Journal of Integrated Design and Process Science*, Preprint, 1–18. https://doi.org/10.3233/JID-220013

Chowdhury, S., & Nath, A. (2001). Trends in natural language processing: Scope and challenges. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, *7*(6), 393–401. https://doi.org/10.32628/CSEIT217698

Cormen, T., Leiserson, C., Rivest, R., & Stein, C. (2009). *Introduction to algorithms* (3rd ed.). MIT Press.

Cornelius, G. W., Doerte, H., Farina, K., & Joachim, R. (2022). *Perception and knowledge of artificial intelligence in healthcare, therapy and diagnostics: A population-representative survey*. medRxiv. https://doi.org/10.1101/2022.12.01.22282960

Cresswell, J. (2014). *Research design. Qualitataive, quantiative and mixed methods approaches* (4th ed.). Sage.

European Parliament. (2016). *Regulation (EU) 2016/679 of the European parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/EC (general data protection regulation)*.

Future of Life Institute. (2023). *Pause giant AI experiments: An open letter*. https://futureoflife.org/open-letter/pause-giant-ai-experiments/

Gal, U. (2023, August 4th). *ChatGPT is a data privacy nightmare. If you've ever posted online, you ought to be concerned*. The Conversation. https://theconversation.com/chatgpt-is-a-data-privacy-nightmare-if-youve-ever-posted-online-you-ought-to-be-concerned-199283

Gilbert, M., & Martin, D. (2022). In search of the moral status of AI: Why sentience is a strong argument. *AI and Society*, *37*, 319–330. https://doi.org/10.1007/s00146-021-01179-z

Greengard, S. (2022, December 29, 2022). *ChatGPT: Understanding the ChatGPT AI chatbot*. eWeek. https://www.eweek.com/big-data-and-analytics/chatgpt/

Halavais, A. (2013). *Search engine society*. Polity Press.

Haque, M., Dharamdasa, I., Sworna, Z., Rajapakse, R., & Ahmad, H. (2022). *"I think this is the most disruptive technology": Exploring sentiments of ChatGPT early adopters using twitter data*. arXiv. https://doi.org/10.48550/arXiv.2212.05856

Hatzius, J., Briggs, J., Kodnani, D., & Pierdomenico, G. (2023). *The potentially large effects of artificial intelligence on economic growth*. Key4biz. https://www.key4biz.it/wp-content/uploads/2023/03/Global-Economics-Analyst_-The-Potentially-Large-Effects-of-Artificial-Intelligence-on-Economic-Growth-Briggs_Kodnani.pdf

International Telecommunication Union. (2021). *Measuring digital development: Facts and figures 2021*. https://www.itu.int/en/ITU-D/Statistics/Pages/facts/default.aspx

Kamalakannan, S., Bhattacharjya, S., Bogdanova, T., Papamitriou, C., Arango-Lasprilla, J., Bentley, J., & Jesus, T. (2021). Health risks and consequences of a COVID-19 infection for people with disabilities: Scoping review and descriptive thematic analysis. *International Journal of Environmental Research and Public Health*, *18*(8), 4348. https://doi.org/10.3390/ijerph18084348

Kanaparthi, V. (2022). Examining Natural Language processing techniques in the education and healthcare fields. *International Journal of Engineering and Advanced Technology*, *12*(2), 8–18. https://doi.org/10.35940/ijeat.B3861.1212222

Kerzel, U. (2020). *Enterprise AI canvas – integrating artificial intelligence into business*. arXiv. https://doi.org/10.48550/arXiv.2009.11190

Lennon, R., Fraleigh, R., Van Scoy, L., Keshaviah, A., Hu, X., Snyder, B., Miller, E., Calo, W., Zgierska, A., & Griffin, C. (2021). Developing and testing an automated qualitative assistant (AQUA) to support qualitative analysis. *Family Medicine and Community Health*, *9*(Suppl 1), e001287. https://doi.org/10.1136/fmch-2021.001287

Macdonald, C., Adeloye, D., Sheikh, A., & Rudan, I. (2023). Can ChatGPT draft a research article? An example of population-level vaccine effectiveness analysis. *Journal of Global Health*, *13*, 01003. https://doi.org/10.7189/jogh.13.01003

McDonnell, A., Jones, M. L., & Read, S. (2000). Practical considerations in case study research: The relationship between methodology and process. *Journal of Advanced Nursing*, *32*(2), 383–390. https://doi.org/10.1046/j.1365-2648.2000.01487.x

Mou, Y., & Xu, K. (2017). The media inequality: Comparing the initial human-human and human-AI social interactions. *Computers in Human Behavior*, *72*, 432–440. https://doi.org/10.1016/j.chb.2017.02.067

Nerdynav. (2023, March 29, 2023). *91 important ChatGPT statistics and facts for March 2023 (Gpt-4, ChatGPT plugins update)*. Nerdy Nav. https://nerdynav.com/chatgpt-statistics/

Ng, G. W., & Leung, W. C. (2020). Strong artificial intelligence and consciousness. *Journal of Artificial Intelligence and Consciousness*, *7*(1), 63–72. https://doi.org/10.1142/S2705078520300042

Oosthuizen, R. (2022). The fourth industrial revolution - Smart technology, artificial intelligence, robotics and algorithms: Industrial psychologists in future workplaces. *Frontiers in Artificial Intelligence*, *5*, 913168. https://doi.org/10.3389/frai.2022.913168

Osadchiy, V., Mills, J. N., & Eleswarapu, S. V. (2020). Understanding patient anxieties in the social media era: Qualitative analysis and Natural Language processing of an online male infertility community. *Journal of Medical Internet Research*, *22*(3), e16728. https://doi.org/10.2196/16728

Popli, N. (2023, March 15, 2023). *GPT-4 has been out for 1 Day. These new projects show just how much more powerful it is*. Time. https://time.com/6263475/gpt4-ai-projects/

Rahman, M., Terano, H., Rahman, N., Salamzadeh, A., & Rahaman, S. (2023). ChatGPT and academic research: A review and recommendations based on practical examples. *Journal of Education, Management and Development Studies in Continuing Education*, *3*(1), 1–12. https://doi.org/10.52631/jemds.v3i1.175

Rogers, E. M. (1962). *Diffusion of innovations*. Free Press.

Russel, S., & Norvig, P. (2016). *Artificial intelligence: A modern approach*. Pearson.

Sanmarchi, F., Bucci, A., & Golinelli, D. (2023). *A step-by-step researcher's guide to the use of an AI- based transformer in epidemiology: An exploratory analysis of ChatGPT using the STROBE checklist for observational studies*. medRxiv. https://doi.org/10.11.01/2023.02.06.23285514

Scharth, M. (2023, February 24, 2023). *How to perfect your prompt writing for ChatGPT, Midjourney and other AI generators*. The Conversation. https://theconversation.com/how-to-perfect-your-prompt-writing-for-chatgpt-midjourney-and-other-ai-generators-198776

Souter, F. (2023, February 25, 2023). *He'd finished a PhD. Then long COVID hit and he couldn't make dinner*. The Age. https://www.theage.com.au/national/he-d-finished-a-phd-then-long-covid-hit-and-he-couldn-t-make-dinner-20230119-p5cdsr.html

Stahl, B. C. (2021). Perspectives on artificial intelligence. In B. C. Stahl (Ed.), *Artificial intelligence for a better future: An ecosystem perspective on the ethics of AI and emerging digital technologies* (pp. 7–17). Springer International Publishing.

Sweeney, A., Greenwood, K. E., Williams, S., Wykes, T., & Rose, D. S. (2013). Hearing the voices of service user researchers in collaborative qualitative data analysis: The case for multiple coding. *Health Expectations*, *16*(4), e89–e99. https://doi.org/10.1111/j.1369-7625.2012.00810.x

van Baalen, S., Boon, M., & Verhoef, P. (2021). From clinical decision support to clinical reasoning support systems. *Journal of Evaluation in Clinical Practice*, *27*(3), 520–528. https://doi.org/10.1111/jep.13541

van Wijngaarden, E., Meide, H. V., & Dahlberg, K. (2017). Researching health care as a meaningful practice: Toward a nondualistic view on evidence for qualitative research. *Qualitative Health Research*, *27*(11), 1738–1747. https://doi.org/10.1177/1049732317711133

Yigitcanlar, T., Degirmenci, K., & Inkinen, T. (2022). Drivers behind the public perception of artificial intelligence: Insights from major Australian cities. *AI and SOCIETY*. https://doi.org/10.1007/s00146-022-01566-0

Zhuo, T., Huang, Y., Chen, C., & Xing, Z. (2023). *Exploring AI ethics of ChatGPT: A diagnostic analysis*. arXiv. https://doi.org/10.48550/arXiv.2301.12867