Education v2
Vol.1 No. 1; December 2024
ISSN: ………………….
URL: https://www.educationv2.com/

Education V2

Eğitim v2
Cilt. 1 Sayı. 1; Aralık 2024
ISSN: ………..………
URL: https://www.educationv2.com/

Qualitative Data Analysis through Generative Artificial Intelligence (GAI)

Seçkin Gök, seckin4501@gmail.com, Mugla Sitki Kocman University, Türkiye,
https://orcid.org/orcid- 0000-0001-6095-9828

SUMMARY

Generative Artificial Intelligence (GAI) is known to have a wide range of applications. In recent years, large language models (LLMs) based on GAI, such as GPT-4, Gemini, and Claude, have also been utilized in qualitative studies, particularly in content analysis processes. These models can analyze large qualitative datasets comprehensively and efficiently within a short time frame. GAI models effectively encode qualitative data and identify categories and themes. While GAI models provide researchers with significant advantages in terms of time and cost efficiency, ethical concerns regarding their use in content analysis have come to the forefront. Moreover, limitations such as biases, the potential to produce misinformation, and superficial analyses are significant barriers to the standalone use of these tools in content analysis. The aim of this study is to comprehensively examine the advantages of using GAI in content analysis and the challenges encountered in this process. In this context, recommendations are provided to researchers planning to utilize GAI models in their analyses, based on a review of relevant literature. Within these recommendations, this study emphasizes that GAI models alone are insufficient for effective content analysis without human involvement and proposes a hybrid approach.

Keywords: Generative artificial intelligence (GAI), large language models (LLMs), qualitative research, content analysis

INTRODUCTION

Recent advancements in artificial intelligence (AI) have demonstrated remarkable exponential growth over the past years. While the foundations of this field were laid in the 1950s with a focus on algorithms and data processing, today, innovations in machine learning and deep learning have enabled AI systems to learn, make decisions, and, in some cases, exhibit characteristics resembling human behavior.

AI is now widely utilized across various domains, including healthcare (Esteva et al., 2017), education (Aktay et al., 2023; Chen et al., 2020), transportation (Litman, 2020), agriculture (Kamilaris & Prenafeta-Boldú, 2018), energy (Kellogg et al., 2018), and disaster management (Bensi et al., 2020). The speed, precision, and efficiency offered by AI tools are among the key reasons for their widespread adoption across these diverse sectors (Gupta et al., 2024). In particular, generative artificial intelligence (GAI), which can produce realistic text, images, and human-like outputs, is currently transforming numerous industries (Bail, 2024). The development of GAI tools has also brought significant changes to academic research, reshaping traditional methodologies and creating new opportunities for data analysis and interpretation (Perkins & Roe, 2024a). A large-scale study conducted by *Nature* indicated that 30% of researchers use GAI-supported language models in their research writing (Prillaman, 2024; cited in Perkins & Roe, 2024b). In this context, it has been suggested that GAI models can also be effectively employed in qualitative research processes such as data coding, theme extraction, and large-scale text analysis (Sivarudran Pillai & Matus, 2024).

Qualitative research is a type of inquiry that seeks to understand and interpret phenomena in their natural settings, focusing on the meanings individuals assign to events (Denzin & Lincoln, 2011, p. 3). In other words, qualitative research involves exploring research problems using interpretive/theoretical frameworks that address the meanings individuals or groups attribute to a social or human issue (Creswell, 2021, p. 45). Data obtained in qualitative research can be analyzed inductively or deductively (Patton, 2018; Ravindran, 2019). Moreover, qualitative research includes designs such as phenomenology, case studies, ethnographic (cultural) studies, and grounded theory (Yıldırım & Şimşek, 2021). Regardless of the research design or approach, the foundation of qualitative research often lies in the meticulous coding process, where researchers identify and label core ideas within qualitative data (Saldaña, 2022). These labels, known as codes, aid researchers in deriving overarching themes and extracting significant insights (Braun & Clarke, 2006, as cited in Gamieldien et al., 2023). In summary, the process of qualitative data analysis can be described as a methodological progression from codes to categories and ultimately to themes (Saldaña, 2022).

Qualitative data often include extensive datasets, such as hundreds of hours of audio recordings and dozens of documents, which can be overwhelming to analyze using traditional methods (e.g., manual analysis) (Katz et al., 2024). Consequently, the analysis of qualitative data requires substantial effort and considerable time (Gamieldien et al., 2023). At this point, GAI tools have been highlighted for their potential to expedite research processes and provide alternative perspectives through AI-generated content (Perkins & Roe, 2024b). Additionally, it has been noted that these tools can support social scientists in conducting content analysis (Patton, 2018) and thematic analysis (Braun & Clarke, 2006), among other commonly used research tasks (Bail, 2024; Perkins & Roe, 2024c). Notably, significant advancements in generative AI tools between November 2022 and June 2023 have captured the attention of both the public and academic communities (Malakar & Leeladharan, 2024), increasing researchers' awareness of these tools.

For decades, qualitative data analysis has relied on software such as ATLAS.ti, Nvivo, Dedoose, and MAXQDA (Katz et al., 2024). With the integration of GAI, these tools have become more functional (Paulus & Marone, 2024). However, recent studies (Bijker et al., 2024; Hamilton et al., 2023; Katz et al., 2024; Lixandru, 2024; Qiao et al., 2024) have shown that GAI tools powered by large language models (LLMs) are increasingly being utilized. Compared to traditional qualitative data analysis software, LLM-based GAI tools are accessible at significantly lower costs. Examples of such tools include OpenAI's ChatGPT, Google's Gemini, Anthropic's Claude, Meta's LLaMA, and Microsoft's Copilot. These GAI tools can perform "zero-shot" coding based on the datasets they are trained on (Davison et al., 2024). However, the use of these LLMs has raised numerous questions and concerns. Their application in qualitative research, particularly for data analysis, has sparked significant debates regarding their appropriateness and ethical implications (Perkins & Roe, 2024a; Sivarudran Pillai & Matus, 2024). Thus, AI presents both new challenges and opportunities for qualitative researchers, solidifying its growing presence in the academic field (Gibson & Beattie, 2024).

In light of these developments, this study examines the use of GAI tools in qualitative data analysis, encompassing both inductive and deductive approaches. The topics discussed below include: the role of LLMs and GAI in research, the application of GAI in qualitative data analysis, challenges and ethical limitations of using GAI, and future perspectives and recommendations.

LLMs and GAI

Natural Language Processing (NLP) is a subfield of AI and linguistics that focuses on creating computer systems capable of processing and analyzing language. NLP underpins many technologies we encounter daily, such as voice assistants (e.g., Siri and Alexa) and translation programs (Nicholas & Bhatia, 2023). The advent of deep learning in the 2010s marked a turning point for NLP. Neural network architectures, including recurrent neural networks and long short-term memory, facilitated the modeling of sequential data. In 2017, the groundbreaking study titled "Attention Is All You Need" (Vaswani et al., 2017) introduced the transformer architecture. This innovation transformed NLP by uncovering relationships within text (Sarraf, 2024).

Specifically, researchers at Google who developed the transformer architecture enabled language models to train on large amounts of data in parallel, rather than sequentially. As a result, transformer-based language models (e.g., ChatGPT) can learn relationships between entire sequences of words rather than processing words individually (Nicholas & Bhatia, 2023). Additionally, these models are capable of processing vast amounts of data simultaneously (Perkins & Roe, 2024b). Such advancements have paved the way for large-scale LLMs capable of handling diverse and complex linguistic tasks. GAI, based on these innovations, has emerged as a natural extension of this development.

GAI refers to machine learning tools trained on large datasets to generate content, such as text, images, or videos, from unstructured data. This technology, rooted in statistics, computer science, and engineering, is also referred to as "foundation models" (Bail, 2024). GAI-powered chatbots go beyond understanding and analyzing language to create new contextual text. Supported by deep learning and large-scale transformers, GAI models produce coherent, contextual, and to some extent, human-like text (Sarraf, 2023, 2024). Due to their training on extensive datasets and the latest technical innovations (e.g., transformer models), these models generate realistic, human-like interactions (Bail, 2024).

For example, OpenAI’s GPT-3 model, introduced in 2020, was trained on half a trillion English words sourced from the internet, while Google’s PaLM model, launched in 2022, was trained on approximately 780 billion words obtained from English websites and social media (Brown et al., 2020; Chowdhery et al., 2022). GAI tools surpass traditional NLP applications in their capabilities. According to Sarraf (2024), these capabilities include:

- **Text generation:** Creating articles, essays, stories, and creative content
- **Language translation:** Performing highly accurate translations between languages
- **Text summarization:** Generating summaries of extensive documents
- **Conversation:** Facilitating human-like dialogue through chatbots and virtual assistants
- **Code generation:** Writing, debugging, and interpreting programming code
- **Sentiment analysis:** Interpreting users' emotions and sentiments from text
- **Personalization:** Providing tailored content for marketing, e-commerce, and education
- **Information retrieval:** Answering questions and extracting information from large datasets
- **Creative problem-solving:** Assisting with brainstorming and generating ideas for creative tasks
- **Domain-specific expertise:** Supporting applications in fields like healthcare, finance, and law

Among GAI-based LLMs, OpenAI's ChatGPT stands out as one of the earliest and most widely recognized tools due to its fluency in language skills and intuitive chatbot-style interaction (Perkins & Roe, 2024b). Following OpenAI’s pioneering efforts, numerous tech giants have rapidly developed GAI-based LLMs. Today, widely used models include Google’s Gemini (formerly Bard), Anthropic’s Claude, Meta's LLaMA, and Microsoft’s Copilot. OpenAI has also released several subsequent models, such as ChatGPT-3.5, GPT-4, GPT-4o, GPT-4o mini and o1-mini. These AI tools have captured public interest and garnered millions of registered users (Curry et al., 2024).

People frequently use GAI tools for various purposes, and researchers find these models particularly functional. Researchers apply GAI tools in data analysis processes for quantitative, qualitative, and mixed-method studies (Combrinck, 2024; Perkins & Roe, 2024a). Notably, the application of LLMs in analyzing qualitative data is gaining attention, as analyses like content analysis and thematic analysis rely on researchers' insights (Sivarudran Pillai & Matus, 2024). In this context, examining how GAI models are used in qualitative research is essential.

The integration of Generative AI (GAI) tools in qualitative data analysis is a rapidly growing trend, offering researchers significant benefits in terms of time, cost, and efficiency. These tools, particularly large language models (LLMs) like ChatGPT and GPT-4, have demonstrated remarkable capabilities in automating various steps of the qualitative analysis process, from coding and categorizing data to identifying themes and insights.

One of the key advantages of using GAI tools in qualitative research is the reduction in time and costs compared to traditional methods. For example, Gilardi et al. (2024) found that ChatGPT was able to analyze 6,183 documents with a level of accuracy that outperformed human workers, while also being significantly cheaper than crowdsourcing platforms like Amazon Mechanical Turk. Similarly, Lixandru (2024) reported that ChatGPT accelerated the analysis process in focus group data, completing it far faster than traditional methods. Gamieldien et al. (2023) highlighted that GPT-3.5, when used to analyze engineering students' exam responses, generated detailed thematic codes comparable to those derived through traditional thematic analysis, outperforming NLP-based cluster-supported methods.

Moreover, some studies have demonstrated the reliability of GAI tools in producing analyses that are in close agreement with human researchers. For instance, Mellon et al. (2024) found that Claude-1.3 closely matched human performance when analyzing survey data, while Hamilton et al. (2023)

identified substantial overlap between the themes found by both human researchers and ChatGPT in semi-structured interviews. The ability of GAI models to efficiently process large datasets has proven particularly useful in research settings where vast amounts of qualitative data need to be analyzed.

However, there are also challenges and limitations. Several studies indicate that while GAI models can perform well in many qualitative analyses, they still struggle with certain aspects of the process. For example, Garg et al. (2024) found that the Cohen's Kappa values, which measure the agreement between the AI and human coding, were not very high (below 0.65), indicating room for improvement in the accuracy of automated coding. Additionally, Curry et al. (2024) noted that while ChatGPT could separate keywords semantically, the categories it produced were often too general and lacked depth, which could hinder more nuanced qualitative analysis.

Despite these limitations, there is a growing consensus that GAI tools have the potential to complement human researchers in the qualitative analysis process. They can serve as valuable secondary coders, enhancing the efficiency and scalability of qualitative research. The development of specialized tools, such as CoAIcoder, PaTaT, and AQUA, demonstrates the growing interest in creating AI-supported solutions for qualitative research (Gao et al., 2023; Gebreegziabher et al., 2023; Lennon et al., 2021).

Furthermore, the use of GAI in qualitative data analysis raises important ethical considerations. Researchers must ensure that the application of these tools adheres to ethical standards, particularly regarding transparency, bias, and accountability in the analysis process (Combrinck, 2024). The increasing reliance on GAI for data analysis highlights the need for responsible use to maintain the integrity and validity of qualitative research.

In conclusion, while GAI tools like ChatGPT and GPT-4 offer substantial benefits for qualitative data analysis, including improved efficiency, cost-effectiveness, and scalability, researchers should be mindful of their limitations and the ethical challenges they pose. As GAI technology continues to evolve, its role in qualitative research is likely to grow, but careful consideration is necessary to ensure that these tools are used in ways that enhance, rather than replace, human insight and judgment in the research process.

The Use of GAI in Qualitative Data Analysis

Analyzing data in qualitative studies is often costly, time-consuming, and prone to errors. Researchers are required to adhere to detailed frameworks during the coding process. GAI tools facilitate a relatively automated approach to coding data (Garg et al., 2024). For instance, Gilardi et al. (2024) analyzed 6,183 documents, including tweets and news articles, using ChatGPT in their qualitative study. They found that ChatGPT accurately categorized the topics of the documents and the authors' perspectives. Moreover, the authors emphasized that ChatGPT outperformed trained human workers and was significantly less costly compared to Amazon Mechanical Turk (MTurk), a crowdsourcing platform developed by Amazon, with a cost as low as $0.003 per task. In another study, Mellon et al. (2024) analyzed survey data from an internet panel on the British election study using both a researcher and six LLMs. The results showed that Claude-1.3 was the most accurate model in approaching the pre-categorized 50 data categories, demonstrating performance close to human accuracy (94%).

In a recent qualitative study, Gamieldien et al. (2023) compared traditional thematic analysis, NLP-based cluster-supported analysis, and GAI-supported analysis using GPT-3.5. The data analyzed consisted of responses from 3,800 engineering students' exam papers (nine questions per paper). The study revealed that the GPT-3.5 model effectively identified themes in student responses similar to those identified through traditional thematic analysis. Furthermore, GPT-3.5 produced more detailed codes compared to NLP-based cluster-supported analysis. These findings suggest that GAI-based language models are more sophisticated and functional than rule-based or keyword-based NLP methods.

Similarly, in a qualitative study by Lixandru (2024), data were obtained from a focus group comprising eight participants and analyzed using ChatGPT. Lixandru highlighted that AI is a reliable tool for interpreting qualitative data and noted that ChatGPT significantly accelerated the analysis process, completing it in a fraction of the time required by traditional methods. In another study, Bijker et al. (2024) employed ChatGPT for content analysis and concluded that ChatGPT holds potential as a secondary coder. However, there are also studies suggesting that GAI tools may not be efficient or sufficient for qualitative data analysis. Garg et al. (2024) conducted a qualitative study involving interviews with 94 middle school teachers about education and educational technologies. The researchers manually categorized the data into eight categories and then analyzed the same data using GPT-3.5-turbo and GPT-4 models. The results showed that the Cohen's Kappa values for both models did not exceed 0.65. Additionally, the GPT-4 model demonstrated better performance compared to GPT-3.5-turbo. Similarly, Curry et al. (2024) utilized ChatGPT for discourse analysis. While ChatGPT performed reasonably well in semantically separating keywords, the categories it generated were overly general and superficial. Hamilton et al. (2023) compared the analyses conducted by researchers and ChatGPT on data obtained from semi-structured interviews with 71 participants. Their findings showed that five themes overlapped between researchers and ChatGPT. While researchers identified six themes in total, ChatGPT generated five themes. Moreover, the analyses conducted by AI lacked the depth, flexibility, and insight demonstrated by human researchers.

In addition to open-source LLMs, some researchers have developed their own AI-supported tools for qualitative data analysis. Gao et al. (2023) developed CoAIcoder; Gebreegziabher et al. (2023) created PaTaT; and Lennon et al. (2021) introduced AQUA for this purpose. From another perspective, Wang (2024) used GPT-3.5, GPT-4.0, and GPT-4o models to conduct semi-structured interviews. Wang's (2024) study demonstrated that GAI models could generate realistic interview responses and offer significant opportunities for improving interview questions.

The use of GAI-based language models among researchers appears to be increasingly widespread. Some researchers argue that GAI models can be utilized for coding data (Siiman et al., 2023), conduct adequate qualitative analyses (Rahman et al., 2023), and perform automated inductive coding (Hämäläinen et al., 2023; Curry et al., 2024). Additionally, qualitative data analyses conducted with these tools have been noted to provide researchers with significant advantages in terms of time and cost savings (Gamieldien et al., 2023; Prescott et al., 2024) as well as efficiency (Pattyn, 2024). Furthermore, it has been emphasized that using GAI tools in data analysis is a natural progression as long as researchers employ these tools responsibly (Combrinck, 2024). However, it is crucial to acknowledge that there are various ethical challenges and considerations related to the use of these tools in the data analysis process.

Challenges and Ethical Limitations in the Use of GAI

The primary limitation to the use of AI tools in qualitative data analysis arises from ethical concerns. In a study reflecting the opinions of 11 researchers on whether GAI should be used for qualitative data analysis (Davison et al., 2024), it was emphasized that a pragmatic perspective does not suffice as a justification for using GAI in such analyses. Instead, the tools employed in research must meet ethical standards and be recognized as ethical by other researchers. Within this framework, Davison et al. (2024) addressed the ethical concerns surrounding the use of GAI in qualitative data analysis under five main categories:

Ethical Concerns

1. **Data Ownership and Rights**

   Sharing data with GAI tools for automated analysis can result in breaches of data privacy. Commercial entities developing these tools may sell the data to third parties, violating ethical standards for research (Bail, 2024).

2. **Data Privacy and Transparency**

   Researchers have a duty to maintain the confidentiality of participants and communities. Using GAI tools may expose sensitive data, raising serious privacy concerns. The lack of transparency in how these tools process and store data exacerbates this issue.

3. **Interpretive Competence**

   Unlike traditional qualitative analysis software, GAI tools independently conduct analyses but lack the ability to interpret nuances of research environments, such as body language, tone of voice, or interpersonal dynamics. GAI tools also struggle to identify indirect language, humor, or cultural subtleties, often leading to findings that are superficial and rigid. This limitation undermines the depth and context essential in qualitative research.

4. **Biases and Discrimination**

   GAI models are often trained on datasets with inherent biases, reflecting societal stereotypes. For example, OpenAI has acknowledged that its models exhibit biases and a Western orientation (Bender et al., 2021). This can lead to discriminatory outcomes and skewed analyses, particularly in studies involving marginalized groups.

5. **Researcher Responsibilities**

   Researchers bear ethical obligations to ensure the **accuracy and integrity** of their work. While GAI tools can assist in analysis, automated applications without human oversight are considered unethical. Although some examples exist (Pavlik, 2023), GAI tools cannot be listed as authors in academic studies. The ethical issues specific to GAI applications evolve over time, highlighting the need for certain standards. For instance, some academic institutions have developed their own GAI services to meet data privacy requirements. This indicates that GAI's ethical concerns are often application-specific, focusing on privacy rather than general ethical issues. Moreover, the ambiguity in current guidelines on the use of GAI for qualitative coding and data analysis can lead to misuse and unethical practices, potentially compromising the reliability of research findings and negatively impacting the development of scientific knowledge. Such practices may also result in skepticism toward findings among practitioners. Notably, organizations such as *Nature* prohibit the use of GAI for specific purposes. Similarly, Wiley explicitly states in its guidelines that GAI models cannot be listed as authors in research articles.

Practical Challenges

1. **"Hallucinations" and Inaccuracies**

   GAI models are prone to "hallucinations," where they generate factually incorrect or fabricated outputs. This poses a major challenge for content and thematic analysis, as findings may include information absent from the dataset (Perkins & Roe, 2024c). Moreover, GAI tools may produce inconsistent results when the same data is reanalyzed, raising concerns about reproducibility and transferability (Lincoln & Guba, 1985).

2. **Limited Understanding of Context**

   Qualitative research often involves complex expressions, including sarcasm, irony, emotion, or culturally specific contexts, which GAI tools struggle to interpret accurately (dos Anjos et al., 2024). These nuances are critical to understanding participant perspectives holistically and avoiding oversimplifications.

3. **Training Data Limitations**

   GAI models are trained on vast datasets sourced from the internet, including social media and websites, which are prone to inaccuracies and misinformation (Feuerriegel et al., 2024). These limitations increase the risk of cognitive errors in analysis, undermining the validity of research outcomes.

4. **Environmental Concerns**

   Training GAI models generates substantial carbon emissions, contributing to environmental degradation (Wang et al., 2023). While some studies argue that GAI tools may produce lower emissions compared to human efforts for the same tasks (Tomlinson et al., 2024), the environmental impact of large-scale AI applications remains a critical concern.

Broader Implications

GAI tools often fail to address the holistic nature of social phenomena, which qualitative research seeks to explore (Karataş, 2015). Participant diversity and perspectives must be handled with care to avoid bias or manipulation, ensuring ethical and accurate representation (Lincoln & Guba, 1985; Patton, 2018). Additionally, the over-reliance on GAI tools risks compromising the interpretive depth of qualitative studies, potentially leading to a loss of trust in research findings.

In conclusion, the ethical and practical challenges associated with GAI tools, including concerns about data privacy, misinformation, bias, and environmental impact, highlight the need for stringent guidelines and human oversight. These issues underscore the importance of balancing innovation with responsibility to ensure that GAI tools are used ethically and effectively in qualitative research.

Future Perspectives and Recommendations

GAI models can automate processes such as text generation, data collection, and data analysis. Additionally, GAI models can identify patterns and trends in data and formulate hypotheses (Dahal, 2024). However, researchers should not perceive GAI tools as substitutes for human judgment and expertise. Particularly, tasks such as interpreting results, developing theoretical frameworks, and drawing inferences should remain the responsibility of researchers. GAI tools should be viewed as instruments that facilitate, support, and accelerate analytical processes rather than replacing critical thinking and domain expertise (Bearman & Ajjawi, 2023; Perkins & Roe, 2024b). For example, while GAI tools may expedite the analysis process, they fall short in interpreting findings, contextualizing data, and conducting in-depth analyses (Perkins & Roe, 2024c). Experienced researchers often outperform GAI in identifying themes (Prescott et al., 2024).

In light of these considerations, a hybrid approach appears necessary for the analysis of qualitative data. Several studies (Combrinck, 2024; Gamieldien et al., 2023; Perkins & Roe, 2024a) emphasize the importance of adopting an integrative hybrid approach to qualitative data analysis. This approach involves synthesizing the capabilities of GAI models with researchers' skills. It can reduce the costs of research, expedite the process, and improve inter-coder reliability (Pattyn, 2024).

The hybrid use of GAI in qualitative data analysis can take various forms. For instance, after GAI performs the initial coding of textual data, analysts can refine these codes to develop more abstract and interpretive themes or adjust existing ones (Bijker et al., 2024; Yan et al., 2024). This method provides researchers with a detailed starting point for analysis (Perkins & Roe, 2024b). Similarly, a researcher could conduct a full thematic analysis and then employ GAI as a secondary analyst to evaluate inconsistencies between codes and themes. This triangulation process (Perkins & Roe, 2024a; Prescott et al., 2024) enhances the credibility of the study, potentially eliminating the need for a second human analyst and adding flexibility to the analytical process (dos Anjos et al., 2024; Prescott et al., 2024). Another approach involves integrating codes generated by researchers with those produced by GAI models during content analysis (Prescott et al., 2024). Such integration allows for leveraging the strengths of both analytical methods in research.

Researchers must adhere to several considerations when using GAI models for qualitative data analysis. They should first be aware of the established ethical principles regarding AI use. According to Skorburg et al. (2024), these principles include taking responsibility for how AI tools are used, safeguarding participants' data privacy, ensuring transparency in AI usage, utilizing AI tools for specific purposes during research, and reporting such usage, as well as mitigating biases inherent in AI systems. Researchers should disclose the source codes and parameters of the AI tools they use and clarify where, how, and for what purpose these tools were employed (Bail, 2024; Dahal, 2024). Furthermore, informed consent forms provided to participants must explicitly state that their data will be shared with AI tools. Additionally, while GAI may provide an initial coding and identify patterns, researchers must thoroughly verify the accuracy of these patterns. They should avoid immediately building their analysis upon preliminary AI-generated codings (Davison et al., 2024). Ultimately, researchers must be cognizant of the errors or biases generated by GAI models and explicitly report the steps they took to validate the AI-generated findings. This can be achieved by cross-checking AI findings with manual analyses or employing multiple GAI tools for triangulation (Perkins & Roe, 2024b). Lastly, researchers should train the analytical tool with a sufficient number of manually coded examples before initiating analysis with GAI models (Garg et al., 2024).

Alan Turing was among the first researchers to propose evaluating AI systems by determining whether humans could distinguish AI-generated content from human-generated content (Bail, 2024). Early attempts to create chatbots capable of passing this test (e.g., ELIZA) were unsuccessful. However, recent studies indicate a shift in this trend. In Mitchell's (2024) study, human participants engaged in five-minute conversations with either a human or various AI tools (ELIZA, GPT-3.5, and GPT-4). The results revealed that GPT-4 was perceived as human 54% of the time, surpassing the 50% threshold and becoming the first LLM to pass the test. Consequently, the use of up-to-date GAI models is crucial for qualitative research, particularly for identifying human-specific contexts and generating codes, categories, and themes (Combrinck, 2024).

CONCLUSION

The development of GAI tools has significantly impacted scientific research, as in many other fields. Researchers strive to incorporate advancements in AI into their workflows. However, caution is paramount when employing GAI tools in research processes. Writing a research text, partially or entirely, through AI tools is unethical and detectable by numerous plagiarism detection tools (e.g., ZeroGPT). Regarding the use of GAI in data analysis, researchers must exercise diligence. While the ability to code vast amounts of data and derive themes within seconds is appealing, GAI models have yet to fully grasp human-specific nuances. Qualitative studies aim to uncover human perspectives in depth, and the superficial and mechanical analyses of GAI tools underscore their limited contribution to qualitative data analysis. The inability of GAI-based LLMs to match researchers' analytical capabilities suggests that these tools will not be suitable for direct (from scratch) use in traditional content analysis in the near future. Nevertheless, these tools appear effective in minimizing human errors and reducing research costs. Therefore, researchers can integrate GAI models—particularly their latest versions—into their studies through a hybrid approach.

REFERENCES

Aktay, S., Gök, S., & Uzunoğlu, D. (2023). ChatGPT in education. *Türk Akademik Yayınlar Dergisi (TAY Journal)*, *7*(2), 378-406. https://doi.org/10.29329/tayjournal.2023.543.03

Alexandra F. Gibson & Alexander Beattie (2024) More or less than human? Evaluating the role of AI-as-participant in online qualitative research. *Qualitative Research in Psychology*, *21*(2), 175-199, https://doi.org/10.1080/14780887.2024.2311427

Bail, C. A. (2024). Can Generative AI improve social science?. *Proceedings of the National Academy of Sciences*, *121*(21), e2314021121.

Bearman, M., & Ajjawi, R. (2023). Learning to work with the black box: Pedagogy for a world with artificial intelligence. *British Journal of Educational Technology*, *54*(5), 1160-1173. https://doi.org/10.1111/bjet.13337

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the dangers of stochastic parrots: Can language models be too big?. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610-623).

Bensi, M. T., Marafi, N. A., & Hoffman, F. O. (2020). Machine learning applications in risk analysis: A review. *Reliability Engineering & System Safety*, *193*, 106596. https://doi.org/10.1016/j.ress.2019.106596

Bijker, R., Merkouris, S. S., Dowling, N. A., & Rodda, S. N. (2024). ChatGPT for Automated Qualitative Research: Content Analysis. *Journal of Medical Internet Research*, *26*(1), e59050. https://doi.org/10.2196/59050

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., … Amodei, D. (2020). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems*, *33*, 1877–1901.

Chen, X., Xie, H., Zou, D., Hwang, G. J., & Li, Y. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Educational Technology & Society*, *23*(4), 1–15.

Chen, Y., Liu, T. X., Shan, Y., & Zhong, S. (2023). The emergence of economic rationality of GPT. *Proceedings of the National Academy of Sciences*, *120*(51), e2316205120.

Chowdhery, A., Narang, S., Devlin, J., Bosma, M., Mishra, G., Roberts, A., Barham, P., Chung, H. W., Sutton, C., Gehrmann, S., Schuh, P., Shi, K., Tsvyashchenko, S., Maynez, J., Rao, A., Barnes, P., Tay, Y., Shazeer, N., Prabhakaran, V., … Fiedel, N. (2023). Palm: Scaling language modeling with pathways. *Journal of Machine Learning Research*, *24*(240), 1-113.

Combrinck, C. (2024). A tutorial for integrating generative AI in mixed methods data analysis. *Discover Education*, *3*(116), 1-23. https://doi.org/10.1007/s44217-024-00214-7

Creswell, J.W. (2021). *Qualitative Inquiry And Research Design: Choosing Among Five Approaches* (M. Bütün & S. B. Demir Trans.). Siyasal Kitabevi. (Orijinal work published 2007).

Curry, N., Baker, P., & Brookes, G. (2024). Generative AI for corpus approaches to discourse studies: A critical evaluation of ChatGPT. *Applied Corpus Linguistics*, *4*(1), 100082. https://doi.org/10.1016/j.acorp.2023.100082

Dahal, N. (2024). How Can Generative AI (GenAI) enhance or hinder qualitative studies? A critical appraisal from south asia, Nepal. *The Qualitative Report*, *29*(3), 722-733.

Davison, R. M., Chughtai, H., Nielsen, P., Marabelli, M., Iannacci, F., van Offenbeek, M., Tarafdar, M., Trenz, M., Techatassanasoontorn, A. A., Andrade, A. D., Panteli, N., & Panteli, N. (2024). The ethics of using generative AI for qualitative data analysis. . *Inf Syst J 34*(5),1433–1439.

Denzin, N. K., & Lincoln, Y. S. (2011). *The Sage handbook of qualitative research*. Sage.

dos Anjos, J. R., de Souza, M. G., de Andrade Neto, A. S., & de Souza, B. C. (2024). An analysis of the generative AI use as analyst in qualitative research in science education. *Revista Pesquisa Qualitativa*, *12*(30), 01-29. https://doi.org/10.33361/RPQ.2024.v.12.n.30.724

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115–118. https://doi.org/10.1038/nature21056

Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). Generative ai. *Business & Information Systems Engineering*, *66*(1), 111-126.

Gamieldien, Y., Case, J. M., & Katz, A. (2023). Advancing qualitative analysis: An exploration of the potential of generative AI and NLP in thematic coding. *Available at SSRN 4487768*.

Gao, J., Choo, K. T. W., Cao, J., Lee, R. K.-W., & Perrault, S. (2023). CoAIcoder: Examining the Effectiveness of AI-assisted Human-to-Human Collaboration in Qualitative Analysis. *ACM Transactions on Computer-Human Interaction*, *31*(1), 1-38. https://doi.org/10.1145/3617362

Garg, R., Han, J., Cheng, Y., Fang, Z., & Swiecki, Z. (2024, March). Automated Discourse Analysis via Generative Artificial Intelligence. In *Proceedings of the 14th Learning Analytics and Knowledge Conference* (pp. 814-820).

Gebreegziabher, S. A., Zhang, Z., Tang, X., Meng, Y., Glassman, E. L., & Li, T. J. J. (2023). PaTAT: HumanAI collaborative qualitative coding with explainable ınteractive rule synthesis. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (pp. 1–19). https://doi.org/10.1145/3544548.3581352

Gilardi, F., Alizadeh, M., & Kubli, M. (2023). ChatGPT outperforms crowd workers for text-annotation tasks. *Proceedings of the National Academy of Sciences*, *120*(30), e2305016120.

Gupta, P., Ding, B., Guan, C., & Ding, D. (2024). Generative AI: A systematic review using topic modelling techniques. *Data and Information Management*, *8*(2), 1-66. https://doi.org/10.1016/j.dim.2024.100066

Hämäläinen, P., Tavast, M., Kunnari, A., 2023. Evaluating large language models in generating synthetic hci research data: a case study. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery* (pp. 1–19). https://doi.org/10.1145/3544548.3580688.

Hamilton, L., Elliott, D., Quick, A., Smith, S., & Choplin, V. (2023). Exploring the use of AI in qualitative analysis: A comparative study of guaranteed income data. *International Journal of Qualitative Methods*, *22*, 16094069231201504. https://doi.org/10.1177/16094069231201504

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147, 70–90. https://doi.org/10.1016/j.compag.2018.02.016

Karataş, Z. (2015). Sosyal bilimlerde nitel araştirma yöntemleri. *Manevi Temelli Sosyal Hizmet Araştırmaları Dergisi*, *1*(1), 62-80.

Katz, A., Fleming, G. C., & Main, J. (2024). Thematic analysis with open-source generative AI and machine learning: A new method for inductive qualitative codebook development. *arXiv preprint arXiv:2410.03721*.

Kellogg, D., Azarian, M., & Pecht, M. (2018). The application of artificial intelligence in the renewable energy sector. *Renewable Energy*, 132, 244–255. https://doi.org/10.1016/j.renene.2018.08.064

Lennon, R. P., Fraleigh, R., Van Scoy, L. J., Keshaviah, A., Hu, X. C., Snyder, B. L., Miller, E. L., Calo, W. A., Zgierska, A. E., & Griffin, C. (2021). Developing and testing an automated qualitative assistant (AQUA) to support qualitative analysis. *Family Medicine and Community Health*, *9*(1), e001287. https://doi.org/10.1136/fmch-2021-001287

Lincoln, Y. S., & Guba, E.G. (1985). *Naturalistic inquiry*. Sage.

Litman, T. (2020). Autonomous vehicle implementation predictions: Implications for transport planning. *Victoria Transport Policy Institute*. https://www.vtpi.org/avip.pdf

Lixandru, D. (2024). The Use of Artificial intelligence for qualitative data analysis: ChatGPT. *Informatica Economica*, *28*(1), 57-67.

Malakar, P., & Leeladharan, M. (2024). Generative AI tools for collaborative content creation: A comparative analysis. *DESIDOC Journal of Library & Information Technology*, *44*(3), 151-157. https://doi.org/10.14429/djlit.44.03.19698

Mellon, J., Bailey, J., Scott, R., Breckwoldt, J., Miori, M., & Schmedeman, P. (2024). Do AIs know what the most important issue is? Using language models to code open-text social survey responses at scale. *Research & Politics*, *11*(1), 1-7. 20531680241231468.

Mitchell, M. (2024). The Turing Test and our shifting conceptions of intelligence. *Science*, *385*(6710), eadq9356.

Nicholas, G. and Bhatia, A. (2023). *Lost in Translation: Large Language Models in Non-English Content Analysis.* Center for Democracy & Technology. https://cdt.org/insights/lostin-translation-large-language-models-in-non-english-content-analysis/

Patton, M. Q. (2018). *Qualitative research & evaluation methods: Integrating theory and practice*

(Bütün, M & S. B. Demir trans.). Pegem Akademi. (Orijinal work published 2014).

Pattyn, F. (2024). The value of generative AI for qualitative research: A pilot study. *Journal of Data Science and Intelligent Systems*, 00(00), 1–8. https://doi.org/10.47852/bonviewJDSIS4202964

Paulus, T. M., & Marone, V. (2024). "In Minutes Instead of Weeks": Discursive Constructions of Generative AI and Qualitative Data Analysis. *Qualitative Inquiry*, 10778004241250065. https://doi.org/10.1177/10778004241250065

Pavlik, J. V. (2023). Collaborating with ChatGPT: Considering the implications of generative artificial intelligence for journalism and media education. *Journalism & Mass Communication Educator*, *78*(1), 84-93. https://doi.org/10.1177/10776958221149577

Perkins, M., & Roe, J. (2024a). Academic publisher guidelines on AI usage: A ChatGPT supported thematic analysis. *F1000Research*, *12*, 1398.

Perkins, M., & Roe, J. (2024b). Generative AI Tools in Academic Research: Applications and Implications for Qualitative and Quantitative Research Methodologies. *arXiv preprint arXiv:2408.06872*.

Perkins, M., & Roe, J. (2024c). The use of Generative AI in qualitative analysis: Inductive thematic analysis with ChatGPT. *Journal of Applied Learning and Teaching*, *7*(1), 390-395.

Prescott, M. R., Yeager, S., Ham, L., Rivera Saldana, C. D., Serrano, V., Narez, J., ... & Montoya, J. (2024). Comparing the efficacy and efficiency of human and generative AI: Qualitative thematic analyses. *JMIR AI*, *3*, e54482. https://doi.org/10.2196/54482

Qiao, S., Fang, X., Garrett, C., Zhang, R., Li, X., & Kang, Y. (2024). Generative AI for Qualitative Analysis in a Maternal Health Study: Coding In-depth Interviews using Large Language Models (LLMs). *medRxiv*, 2024-09.

Rahman, M. M., Terano, H. J., Rahman, M. N., Salamzadeh, A., & Rahaman, M. S. (2023). ChatGPT and academic research: A review and recommendations based on practical examples. ChatGPT and Academic Research: A Review and Recommendations Based on Practical Examples. *Journal of Education, Management and Development Studies*, *3*(1), 1-12. https://doi.org/10.52631/jemds.v3i1.175

Ravindran, V. (2019). Data analysis in qualitative research. *Indian Journal of Continuing Nursing Education*, *20*(1), 40-45. https://doi.org/10.4103/IJCN.IJCN_1_19

Saldana, J. (2019). *The coding manual for qualitative researchers.* (A. Tüfekçi & N. Şad trans.). Pegem. (Orijinal work published 2015).

Sarraf, S. (2024). Evaluating generative AI-enhanced content: A conceptual framework using qualitative, quantitative, and mixed-methods approaches. *arXiv preprint arXiv:2411.17943*.

Sarraf, S. H., Rastegari, S., & Soltanieh, M. (2023). Deposition of mono dispersed Co–CeO2 nanocomposite coatings by a sol-enhanced pulsed reverse electroplating: process parameters screening. *Journal of Materials Research and Technology*, *23*, 3772-3789.

Siiman, L. A., Rannastu-Avalos, M., Pöysä-Tarhonen, J., Häkkinen, P., & Pedaste, M. (2023, August). Opportunities and challenges for AI-assisted qualitative data analysis: An example from collaborative problem-solving discourse data. In *International Conference on Innovative Technologies and Learning* (pp. 87-96). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-40113-89

Sivarudran Pillai, V., & Matus, K. (2024). Regulatory solutions to alleviate the risks of generative AI models in qualitative research. *Journal of Asian Public Policy*, 1-24. (Published online: 08 Sep 2024). https://doi.org/10.1080/17516234.2024.2399098

Skorburg J.A., O'Doherty, K., & Friesen, P. (2024). Persons or data points? Ethics, artifcial intelligence, and the participatory turn in mental health research. *American Psychologis*, *79*(1), 137-149. https://doi.org/10.1037/amp0001168.

Tomlinson, B., Black, R. W., Patterson, D. J., & Torrance, A. W. (2024). The carbon emissions of writing and illustrating are lower for AI than for humans. *Scientific Reports*, *14*(1), 3732.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017, December 5). Attention is all you need. *Advances in Neural Information Processing Systems*. [perma. cc/2ZDX-Z796].

Wang, X., Na, C., Strubell, E., Friedler, S., & Luccioni, S. (2023). Energy and Carbon Considerations of Fine-Tuning BERT. *arXiv preprint arXiv:2311.10267*.

Wang, Y. (2024). Interviewing GPTs: Can generative AI intelligence become a new source of qualitative research data?. *Available at SSRN 4874356*.

Yan, L., Echeverria, V., Fernandez-Nieto, G. M., Jin, Y., Swiecki, Z., Zhao, L., Gašević, D., & MartinezMaldonado, R. (2024). Human-AI collaboration in thematic analysis using ChatGPT: A user study and design recommendations. In *Extended Abstracts of the 2024 CHI Conference on Human Factors in Computing Systems* (pp. 1–7). https://doi.org/10.1145/3613905.3650732

Yıldırım, A., & Şimşek, H. (2021). *Sosyal bilimlerde nitel araştırma yöntemleri.* Seçkin Yayıncılık.