

# AI and Large Language Models in Qualitative Data Analysis: Current Applications in Social Sciences and HCI Research

## 1. Introduction

The integration of Artificial Intelligence (AI), particularly Large Language Models (LLMs) such as GPT-3, GPT-4, and their open-source counterparts, is rapidly transforming qualitative data analysis (QDA) in the social sciences and Human-Computer Interaction (HCI) research. LLMs are now being used to automate or augment a range of qualitative tasks, including inductive and deductive coding, thematic analysis, content summarization, and even the simulation of interview participants (De Paoli, 2023; Xiao et al., 2023; Qiao et al., 2025; Player et al., 2025; Zhang et al., 2023; Hayes, 2025; Paoli, 2023; Mathis et al., 2024; Tai et al., 2024; Zhang et al., 2024). Studies demonstrate that LLMs can significantly accelerate the coding process, reduce costs, and handle large datasets with high efficiency, often achieving moderate to substantial agreement with human coders (De Paoli, 2023; Qiao et al., 2025; Player et al., 2025; Mathis et al., 2024; Tai et al., 2024; Qiao et al., 2024; Li et al., 2024). Researchers are leveraging LLMs for both traditional qualitative workflows (e.g., thematic analysis, grounded theory) and novel applications such as AI-augmented surveys, conversational analysis, and the development of new digital research tools (Kantor, 2023; Karjus, 2023; Fischer & Biemann, 2024; Rasheed et al., 2024; Hayes, 2025; Ke & Ng, 2024; MacGeorge, 2025). However, the literature also highlights important limitations: LLMs may struggle with context, nuance, and cultural specificity, can introduce systematic biases, and require careful prompt engineering and human oversight to ensure validity and ethical integrity (Zhang et al., 2023; Roberts et al., 2024; Amirova et al., 2023; Ashwin et al., 2023; Sakaguchi et al., 2025; Friedman et al., 2024; Dengel et al., 2023; Cox et al., 2024). The consensus is that LLMs are best used as collaborative partners—augmenting, rather than replacing, human expertise in qualitative research (Crocker et al., 2025; Kirsten et al., 2025; Bano et al., 2024; Jalali & Akhavan, 2024; Namvarpour & Razi, 2024; Hayes, 2025; Thominet et al., 2024; Zhang et al., 2024; Bano et al., 2023; Li et al., 2024). This review synthesizes the state of the art, best practices, and ongoing challenges in the use of AI and LLMs for qualitative data analysis in social sciences and HCI.

## 2. Methods

We conducted a comprehensive search across over 170 million research papers in Consensus, including Semantic Scholar, PubMed, and other major databases. The search targeted the use of AI and LLMs in qualitative data analysis within social sciences and HCI. In total, 1,022 papers were identified, 752 were screened, 573 were deemed eligible, and the 50 most relevant papers were included in this review.

## Search Strategy



FIGURE 1 Flow diagram of the literature search and selection process.

Twenty unique searches were executed across eight search groups, systematically covering foundational perspectives, methodological diversity, critical and interdisciplinary views, and human-AI collaboration in qualitative research.

## 3. Results

### 3.1. Core Applications of LLMs in Qualitative Data Analysis

LLMs are being used for both **inductive and deductive coding**, thematic analysis, and content summarization. Studies show that LLMs can infer main themes from interview data, often matching or complementing human-coded results (De Paoli, 2023; Xiao et al., 2023; Qiao et al., 2025; Player et al., 2025; Paoli, 2023; Mathis et al., 2024; Tai et al., 2024; Zhang et al., 2024; Qiao et al., 2024; Li et al., 2024). LLMs are also used for document classification, information extraction, and generating representative quotes or summaries (Fischer & Biemann, 2024; Player et al., 2025; Hayes, 2025; Mathis et al., 2024; Zhang et al., 2024). In HCI, LLMs are integrated into QDA platforms as opt-in assistants, supporting tasks such as data preprocessing, onboarding, and mediation (Kirsten et al., 2025; Fischer & Biemann, 2024; Hayes, 2025; Ke & Ng, 2024).

### 3.2. Human-AI Collaboration and Workflow Integration

The literature emphasizes a **human-in-the-loop** approach, where LLMs augment but do not replace human expertise (Crocker et al., 2025; Kirsten et al., 2025; Bano et al., 2024; Jalali & Akhavan, 2024; Namvarpour & Razi, 2024; Hayes, 2025; Thominet et al., 2024; Zhang et al., 2024; Bano et al., 2023; Li et al., 2024). Researchers use LLMs to accelerate initial coding, surface patterns, and suggest themes, while retaining responsibility for prompt design, interpretation, and contextualization (Zhang et al., 2023; Hayes, 2025; Thominet et al., 2024; Zhang et al., 2024). Frameworks have been developed to guide responsible integration, spanning minimal to high AI involvement, and highlighting the need for transparency, prompt engineering, and iterative validation (Kirsten et al., 2025; Zhang et al., 2023; Hayes, 2025; Thominet et al., 2024).

### 3.3. Performance, Efficiency, and Validation

LLMs can dramatically reduce the time and cost of qualitative analysis—sometimes by over 80%—while maintaining moderate to substantial agreement with human coders (Cohen’s Kappa often >0.7 for descriptive themes) (Qiao et al., 2025; Player et al., 2025; Mathis et al., 2024; Tai et al., 2024; Qiao et al., 2024; Li et al., 2024). However, performance varies by task complexity, language, and the need for cultural or emotional nuance (Qiao et al., 2025; Zhang et al., 2023; Mathis et al., 2024; Sakaguchi et al., 2025). LLMs excel at identifying descriptive themes but may miss deeper, context-dependent, or culturally embedded meanings (Zhang et al., 2023; Sakaguchi et al., 2025; Dengel et al., 2023).

### 3.4. Limitations, Risks, and Ethical Considerations

Key limitations include **loss of nuance**, context insensitivity, systematic bias, hallucination, and challenges in handling non-English or culturally complex data (Zhang et al., 2023; Roberts et al., 2024; Amirova et al., 2023; Ashwin et al., 2023; Sakaguchi et al., 2025; Friedman et al., 2024; Dengel et al., 2023; Cox et al., 2024). LLMs may introduce non-random errors, especially with marginalized or underrepresented groups (Ashwin et al., 2023; Friedman et al., 2024; Dengel et al., 2023). Ethical concerns include data privacy, transparency, and the risk of over-reliance on AI outputs (Zhang et al., 2023; Roberts et al., 2024; Ashwin et al., 2023; Friedman et al., 2024; Dengel et al., 2023; Cox et al., 2024). Best practices recommend human oversight, transparent reporting, and iterative validation to mitigate these risks (Kantor, 2023; Zhang et al., 2023; Roberts et al., 2024; Hayes, 2025; Abdurahman et al., 2025; Ashwin et al., 2023; Friedman et al., 2024; Dengel et al., 2023; Li et al., 2024).

#### Key Papers

Paper	Application/Task	Methodology	Key Results	Limitations
(De Paoli, 2023)	Inductive thematic analysis	LLM (GPT-3.5) vs. human	LLM inferred most main themes; viable for TA	Partial reproduction of human nuance
(Xiao et al., 2023)	Deductive coding	GPT-3 + codebook	Fair to substantial agreement with experts	Generalizability, prompt design
(Player et al., 2025)	Automated thematic analysis	DECOTA (LLM+STM)	91.6% code agreement, 378x faster than humans	Needs human interpretation for context
(Mathis et al., 2024)	Thematic analysis (healthcare)	Open-source LLM vs. human	Moderate to substantial similarity (Jaccard 0.44–0.69)	Cultural/contextual limitations
(Li et al., 2024)	Health interview analysis	GPT-4 vs. human	Moderate agreement ( $\kappa=0.401$ ); AI consistent for key themes	Missed subthemes, less interpretive depth

**FIGURE 2** Comparison of key studies on LLM-assisted qualitative data analysis.

## Top Contributors







Type	Name	Papers
Author	He Zhang	(Zhang et al., 2023; Zhang et al., 2024)
Author	Shan Qiao	(Qiao et al., 2025; Qiao et al., 2024)
Author	S. Paoli	(De Paoli, 2023; Paoli, 2023; Mathis et al., 2024)
Journal	<i>Journal of Medical Internet Research</i>	(Sakaguchi et al., 2025; Yue et al., 2025; Li et al., 2024)
Journal	<i>International Journal of Qualitative Methods</i>	(Hayes, 2025; Tai et al., 2024; Parker et al., 2023)
Journal	<i>ArXiv</i>	(Kirsten et al., 2025; Breazu et al., 2024; Jalali & Akhavan, 2024; Namvarpour & Razi, 2024; Torii et al., 2024; Paoli, 2023; Thominet et al., 2024; Zhang et al., 2024; MacGeorge, 2025)

**FIGURE 3** Authors & journals that appeared most frequently in the included papers.

## 4. Discussion

The use of AI and LLMs in qualitative data analysis is rapidly evolving, offering substantial efficiency gains and new methodological possibilities for social sciences and HCI research (De Paoli, 2023; Xiao et al., 2023; Qiao et al., 2025; Player et al., 2025; Zhang et al., 2023; Hayes, 2025; Paoli, 2023; Mathis et al., 2024; Tai et al., 2024; Zhang et al., 2024; Qiao et al., 2024; Li et al., 2024). LLMs can reliably assist with coding, theme identification, and summarization, especially for large datasets and descriptive tasks (De Paoli, 2023; Xiao et al., 2023; Qiao et al., 2025; Player et al., 2025; Mathis et al., 2024; Tai et al., 2024; Qiao et al., 2024; Li et al., 2024). However, their limitations—particularly in handling context, nuance, and cultural specificity—necessitate a collaborative, human-in-the-loop approach (Zhang et al., 2023; Roberts et al., 2024; Hayes, 2025; Thominet et al., 2024; Ashwin et al., 2023; Sakaguchi et al., 2025; Friedman et al., 2024; Dengel et al., 2023; Cox et al., 2024). The literature underscores the importance of prompt engineering, transparent reporting, and iterative validation to ensure quality and ethical integrity (Kantor, 2023; Zhang et al., 2023; Roberts et al., 2024; Hayes, 2025; Abdurahman et al., 2025; Ashwin et al., 2023; Friedman et al., 2024; Dengel et al., 2023; Li et al., 2024). As LLMs continue to advance, their role in qualitative research will likely expand, but human expertise remains essential for interpretation, contextualization, and ethical oversight (Crocker et al., 2025; Kirsten et al., 2025; Bano et al., 2024; Jalali & Akhavan, 2024; Namvarpour & Razi, 2024; Hayes, 2025; Thominet et al., 2024; Zhang et al., 2024; Bano et al., 2023; Li et al., 2024).

## Claims and Evidence Table

Claim	Evidence Strength	Reasoning	Papers
LLMs can accelerate and scale qualitative coding and thematic analysis	 Strong	Multiple studies show dramatic time/cost savings and moderate to substantial agreement with human coders	(De Paoli, 2023; Xiao et al., 2023; Qiao et al., 2025; Player et al., 2025; Mathis et al., 2024; Tai et al., 2024; Zhang et al., 2024; Qiao et al., 2024; Li et al., 2024)
Human-in-the-loop approaches yield the best results, combining AI efficiency with human nuance	 Strong	Consensus across empirical and framework papers; AI augments, not replaces, human expertise	(Crocker et al., 2025; Kirsten et al., 2025; Bano et al., 2024; Jalali & Akhavan, 2024; Namvarpour & Razi, 2024; Hayes, 2025; Thominet et al., 2024; Zhang et al., 2024; Bano et al., 2023; Li et al., 2024)
LLMs struggle with context, nuance, and cultural specificity	 Moderate	Studies show lower agreement for complex, context-dependent, or non-English data	(Zhang et al., 2023; Roberts et al., 2024; Amirova et al., 2023; Mathis et al., 2024; Ashwin et al., 2023; Sakaguchi et al., 2025; Friedman et al., 2024; Dengel et al., 2023; Cox et al., 2024)
Prompt engineering and validation are critical for reliable AI outputs	 Moderate	Performance depends on prompt design and iterative human validation	(Xiao et al., 2023; Zhang et al., 2023; Hayes, 2025; López-Pérez et al., 2025; Thominet et al., 2024; Frigui, 2024; Zhang et al., 2024)
LLMs can introduce systematic bias and require ethical oversight	 Moderate	Evidence of non-random errors, especially with marginalized groups; privacy and transparency concerns	(Zhang et al., 2023; Roberts et al., 2024; Ashwin et al., 2023; Friedman et al., 2024; Dengel et al., 2023; Cox et al., 2024)
LLMs are less effective for deep interpretive or theory-building tasks	 Moderate	AI misses subthemes, latent meanings, and complex connections	(Zhang et al., 2023; Roberts et al., 2024; Mathis et al., 2024; Sakaguchi et al., 2025; Friedman et al., 2024; Dengel et al., 2023; Li et al., 2024)

**FIGURE 4** Key claims and support evidence identified in these papers.

## 5. Conclusion

AI and Large Language Models are transforming qualitative data analysis in social sciences and HCI, offering unprecedented efficiency and scalability. While LLMs can reliably assist with coding and theme identification, their limitations in context, nuance, and ethics mean that human expertise remains indispensable for high-quality, trustworthy qualitative research.

### 5.1. Research Gaps

Key gaps include improving LLM performance for non-English and culturally complex data, developing robust validation and bias mitigation strategies, and expanding applications to deeper interpretive and theory-building tasks.

#### Research Gaps Matrix

Application/Task	English Data	Non-English Data	Descriptive Coding	Interpretive/Theory-Building	Bias/Ethics
Inductive Coding	8	2	7	3	2
Deductive Coding	7	2	6	2	2
Thematic Analysis	8	2	7	3	2
Summarization/Extraction	7	1	6	2	1
Bias/Ethics	2	1	2	1	8

**FIGURE 5** Heatmap of research coverage by application/task and study attribute.

### 5.2. Open Research Questions

Future research should focus on enhancing LLM interpretability for complex, non-English, and culturally nuanced data, developing robust validation and bias mitigation protocols, and exploring the integration of LLMs in deeper interpretive and theory-building qualitative research.

Question	Why
How can LLMs be improved to handle context, nuance, and cultural specificity in qualitative data analysis?	Addressing this will expand the applicability and trustworthiness of AI-assisted qualitative research across diverse populations.
What are the most effective strategies for validating and mitigating bias in LLM-assisted qualitative analysis?	Ensuring reliability and fairness is critical for ethical and scientific integrity in social research.
How can LLMs be integrated into deeper interpretive and theory-building tasks in qualitative research?	Moving beyond descriptive coding will unlock new methodological possibilities and advance the field.

**FIGURE 6** Open research questions and their significance for future work.

In summary, LLMs are powerful tools for scaling and accelerating qualitative data analysis, but their optimal use requires human expertise, careful validation, and ongoing methodological innovation.

*These papers were sourced and synthesized using Consensus, an AI-powered search engine for research. Try it at <https://consensus.app>*

## References

- De Paoli, S. (2023). Performing an Inductive Thematic Analysis of Semi-Structured Interviews With a Large Language Model: An Exploration and Provocation on the Limits of the Approach. *Social Science Computer Review*, 42, 997 - 1019. <https://doi.org/10.1177/08944393231220483>
- Kantor, J. (2023). Best practices for implementing ChatGPT, large language models, and artificial intelligence in qualitative and survey-based research. *JAAD International*, 14, 22 - 23. <https://doi.org/10.1016/j.jdin.2023.10.001>
- Crocker, A., Schmidt, M., Tejeda, J., & Rodriguez-Mori, H. (2025). Proof of Concept: Leveraging Large Language Models for Qualitative Analysis of Participant Feedback. *Journal of Extension*. <https://doi.org/10.34068/joe.63.01.16>
- Kirsten, E., Buckmann, A., Lassak, L., Borgert, N., Mhaidli, A., & Becker, S. (2025). From Assistance to Autonomy - A Researcher Study on the Potential of AI Support for Qualitative Data Analysis. *ArXiv*, abs/2501.19275. <https://doi.org/10.48550/arXiv.2501.19275>
- Karjus, A. (2023). Machine-assisted quantizing designs: augmenting humanities and social sciences with artificial intelligence. *Humanities and Social Sciences Communications*. <https://doi.org/10.1057/s41599-025-04503-w>
- Fischer, T., & Biemann, C. (2024). Exploring Large Language Models for Qualitative Data Analysis. *Proceedings of the 4th International Conference on Natural Language Processing for Digital Humanities*. <https://doi.org/10.18653/v1/2024.nlp4dh-1.41>
- Xiao, Z., Yuan, X., Liao, Q., Abdelghani, R., & Oudeyer, P. (2023). Supporting Qualitative Analysis with Large Language Models: Combining Codebook with GPT-3 for Deductive Coding. *Companion Proceedings of the 28th International Conference on Intelligent User Interfaces*. <https://doi.org/10.1145/3581754.3584136>

- Qiao, S., Fang, X., Wang, J., Zhang, R., Li, X., & Kang, Y. (2025). Generative AI for thematic analysis in a maternal health study: coding semistructured interviews using large language models. *Applied Psychology. Health and Well-Being*, 17. <https://doi.org/10.1111/aphw.70038>
- Rasheed, Z., Waseem, M., Ahmad, A., Kemell, K., Wang, X., Nguyen-Duc, A., & Abrahamsson, P. (2024). Can Large Language Models Serve as Data Analysts? A Multi-Agent Assisted Approach for Qualitative Data Analysis. *ArXiv*, abs/2402.01386. <https://doi.org/10.48550/arXiv.2402.01386>
- Player, L., Hughes, R., Mitev, K., Whitmarsh, L., Demski, C., Nash, N., Papakonstantinou, T., & Wilson, M. (2025). The use of large language models for qualitative research: The Deep Computational Text Analyser (DECOTA).. *Psychological methods*. <https://doi.org/10.1037/met0000753>
- Breazu, P., Schirmer, M., Hu, S., & Katsos, N. (2024). Large Language Models and Thematic Analysis: Human-AI Synergy in Researching Hate Speech on Social Media. *ArXiv*, abs/2408.05126. <https://doi.org/10.48550/arXiv.2408.05126>
- Zhang, H., Wu, C., Xie, J., Lyu, Y., Cai, J., & Carroll, J. (2023). Redefining Qualitative Analysis in the AI Era: Utilizing ChatGPT for Efficient Thematic Analysis. *ArXiv*, abs/2309.10771. <https://doi.org/10.1016/j.chbah.2025.100144>
- Roberts, J., Baker, M., & Andrew, J. (2024). Artificial intelligence and qualitative research: The promise and perils of large language model (LLM) ‘assistance’. *Critical Perspectives on Accounting*. <https://doi.org/10.1016/j.cpa.2024.102722>
- Bano, M., Zowghi, D., & Whittle, J. (2024). AI and Human Reasoning: Qualitative Research in the Age of Large Language Models. *AI Ethics Journal*. <https://doi.org/10.47289/aiej20240122>
- Jalali, M., & Akhavan, A. (2024). Integrating AI Language Models in Qualitative Research: Replicating Interview Data Analysis with ChatGPT. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4714998>
- Namvarpour, M., & Razi, A. (2024). Apprentices to Research Assistants: Advancing Research with Large Language Models. *ArXiv*, abs/2404.06404. <https://doi.org/10.48550/arXiv.2404.06404>
- Torii, M., Murakami, T., & Ochiai, Y. (2024). Expanding Horizons in HCI Research Through LLM-Driven Qualitative Analysis. *ArXiv*, abs/2401.04138. <https://doi.org/10.48550/arXiv.2401.04138>
- Hayes, A. (2025). “Conversing” With Qualitative Data: Enhancing Qualitative Research Through Large Language Models (LLMs). *International Journal of Qualitative Methods*. <https://doi.org/10.1177/16094069251322346>
- Abdurahman, S., Ziabari, A., Moore, A., Bartels, D., & Dehghani, M. (2025). A Primer for Evaluating Large Language Models in Social-Science Research. *Advances in Methods and Practices in Psychological Science*. <https://doi.org/10.1177/25152459251325174>
- Paoli, S. (2023). Can Large Language Models emulate an inductive Thematic Analysis of semi-structured interviews? An exploration and provocation on the limits of the approach and the model. *ArXiv*, abs/2305.13014. <https://doi.org/10.48550/arXiv.2305.13014>
- Amirova, A., Fteropoulli, T., Ahmed, N., Cowie, M., & Leibo, J. (2023). Framework-based qualitative analysis of free responses of Large Language Models: Algorithmic fidelity. *PLOS ONE*, 19. <https://doi.org/10.1371/journal.pone.0300024>
- Mathis, W., Zhao, S., Pratt, N., Weleff, J., & Paoli, S. (2024). Inductive thematic analysis of healthcare qualitative interviews using open-source large language models: How does it compare to traditional methods?. *Computer methods and programs in biomedicine*, 255, 108356. <https://doi.org/10.1016/j.cmpb.2024.108356>



López-Pérez, B., Chen, Y., Li, X., Cheng, S., & Razavi, P. (2025). Exploring the potential of large language models to understand interpersonal emotion regulation strategies from narratives.. *Emotion*.

<https://doi.org/10.1037/emo0001528>

Tai, R., Bentley, L., Xia, X., Sitt, J., Fankhauser, S., Chicas-Mosier, A., & Monteith, B. (2024). An Examination of the Use of Large Language Models to Aid Analysis of Textual Data. *International Journal of Qualitative Methods*, 23.

<https://doi.org/10.1177/16094069241231168>

Thominet, L., Amorim, J., Acosta, K., & Sohan, V. (2024). Role Play: Conversational Roles as a Framework for Reflexive Practice in AI-Assisted Qualitative Research. *Journal of Technical Writing and Communication*, 54, 396 - 418. <https://doi.org/10.1177/00472816241260044>

Frigui, E. (2024). Leveraging Large Language Models for Automating Inductive Qualitative Coding: A Comparative Study of Prompt Engineering Techniques. *Inquiry@Queen's Undergraduate Research Conference Proceedings*.

<https://doi.org/10.24908/iqurcp18054>

Ashwin, J., Chhabra, A., & Rao, V. (2023). Using Large Language Models for Qualitative Analysis can Introduce Serious Bias. *ArXiv*, abs/2309.17147. <https://doi.org/10.48550/arXiv.2309.17147>

Sakaguchi, K., Sakama, R., & Watari, T. (2025). Evaluating ChatGPT in Qualitative Thematic Analysis With Human Researchers in the Japanese Clinical Context and Its Cultural Interpretation Challenges: Comparative Qualitative Study. *Journal of Medical Internet Research*, 27. <https://doi.org/10.2196/71521>

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

Parker, R., Mancini, K., & Abram, M. (2023). Natural Language Processing Enhanced Qualitative Methods: An Opportunity to Improve Health Outcomes. *International Journal of Qualitative Methods*, 22.

<https://doi.org/10.1177/16094069231214144>

Ke, P., & Ng, K. (2024). Human-AI Synergy in Survey Development: Implications from Large Language Models in Business and Research. *ACM Transactions on Management Information Systems*, 16, 1 - 39.

<https://doi.org/10.1145/3700597>

Yue, Y., Liu, D., Lv, Y., Hao, J., & Cui, P. (2025). A Practical Guide and Assessment on Using ChatGPT to Conduct Grounded Theory: Tutorial. *Journal of Medical Internet Research*, 27. <https://doi.org/10.2196/70122>

Friedman, C., Owen, A., & VanPuymbrouck, L. (2024). Should ChatGPT help with my research? A caution against artificial intelligence in qualitative analysis. *Qualitative Research*. <https://doi.org/10.1177/14687941241297375>

Dengel, A., Gehrlein, R., Fernes, D., Görlich, S., Maurer, J., Pham, H., Großmann, G., & Eisermann, N. (2023). Qualitative Research Methods for Large Language Models: Conducting Semi-Structured Interviews with ChatGPT and BARD on Computer Science Education. *Informatics*, 10, 78. <https://doi.org/10.3390/informatics10040078>

MacGeorge, R. (2025). Conversations With my Data: Exploring the Potential of Large Language Models in Qualitative Futures Research. *World Futures Review*. <https://doi.org/10.1177/19467567251330210>

Qiao, S., Fang, X., Wang, J., Zhang, R., Li, X., & Kang, Y. (2024). Generative AI for Thematic Analysis in a Maternal Health Study: Coding Semi-structured Interviews using Large Language Models (LLMs). *medRxiv*.

<https://doi.org/10.1101/2024.09.16.24313707>

Bano, M., Zowghi, D., & Whittle, J. (2023). Exploring Qualitative Research Using LLMs. *ArXiv*, abs/2306.13298.

<https://doi.org/10.48550/arXiv.2306.13298>

Li, K., Fernandez, A., Schwartz, R., Rios, N., Carlisle, M., Amend, G., Patel, H., & Breyer, B. (2024). Comparing GPT-4 and Human Researchers in Health Care Data Analysis: Qualitative Description Study. *Journal of Medical Internet Research*, 26. <https://doi.org/10.2196/56500>

Cox, E., Shirani, F., & Rouse, P. (2024). Voices from the algorithm: Large language models in social research. *Energy Research & Social Science*. <https://doi.org/10.1016/j.erss.2024.103559>