Generative AI as a Tool for Thematic Analysis: An Exploratory Study with ChatGPT

Sian Joel-Edgar & Yu-Chun Pan
Northeastern University London, London, UK
sian.joel-edgar@nulondon.ac.uk; yu-chun.pan@nulondon.ac.uk

Abstract

Artificial intelligence (AI) has seen rapid development in recent years and it has increasingly applied to various fields. Research is no exception. However, there is much to be explored in this domain. This study aims to explore the suitability of current generative AI applications for research purposes. The focus is on the generative AI's capability to synthesise information as a potential alternative or supplement to human-based information synthesisation. In order to evaluate the effectiveness of the thematic analysis produced by generative AI, this study compares the generative AI-produced results by ChatGPT with human-generated results, based on the same set of papers. The results show generative AI produced very similar results to humans, in terms of the topics themselves and the number of topics identified. However, there are also some minor mismatches between generative AI and human results.

Keywords: AI, Artificial Intelligence, ChatGPT, ChatGPT4, Research Methods, Thematic Analysis

1.0 Introduction

Artificial Intelligence (AI) has made its way into different aspects of our lives and attracted attention from several domains. AI can be considered a system's capability to correctly interpret and learn from external data and to use the learning to achieve specific goals and tasks through flexible adaptation (Makarius et al, 2020). The ability of AI has come a long way since models based on decision trees, random forests and k-means clusterings. Generative AI, a field within artificial intelligence (AI), is responsible for generating fresh and potentially unique content (van Dis et al., 2023). Its application can be viewed as both a creative and rational tool, depending on its usage and the surrounding circumstances. With the capacity of natural language processing powered by supervised and unsupervised training, generative AI applications demonstrated a wide range of use from creative writing to business document generation (Metz, 2023). In November 2022, OpenAI introduced ChatGPT, which swiftly garnered acclaim for its innovative approach to generating AI-based content (Dwivedi et al, 2023). ChatGPT, as one of the most commonly used generative AI applications, provides unique text in response to user queries by harnessing a huge collection of textual data. The outputs closely mimic human-generated content. There has been widespread usage of ChatGPT in a variety of fields, such as software development, poetry, essays, corporate communication, and legal agreements (Zhuo et al, 2023).

AI has significantly impacted societies and individuals. Organisations are implementing AI in their business process at a fast pace (De Cremer & Kaparov, 2021). It has been applied to marketing (Manis, et al., 2023), social media marketing (Liu et al., 2023), engineering design (Yüksel et al., 2023), healthcare (Mahdi et al., 2023), mental health (Thieme et al., 2023), banking (Rahman et al., 2023), human resource management (Chowdhury et al., 2023) and so on. When it comes to research, several research papers have listed ChatGPT as authors (Stokel-Walker, 2023) which caused debates on whether generative AI applications can be considered as credited authors. There is no universally agreed guidance on such stances, and some of the most prestigious academic journals, such as Nature and Science, have made it clear that they do not accept generative AI applications as authors. Whether generative AI applications should be listed as credited authors or not, researchers are increasingly adopting generative AI for research purposes. Therefore, there is a need to advance our understanding of such implementations.

This study aims to explore the suitability of generative AI as a tool for research purposes, thematic analysis in particular. The following literature review will provide an overview of the AI applications and generative AI for research purposes. An experiment that evaluates the effectiveness of a thematic analysis conducted by a generative AI application, namely ChatPGT, will be introduced. The findings will highlight the iterative steps this study takes and the restless, followed discussion and conclusion.

2.0 Relevant Work

With the rapid advancement of AI algorithms, the content generated by AI, such as social media feeds, can be indistinguishable from human-generated content (Rossi et al., 2023). The availability of several AI applications, such as ChatGPT, Bard, Microsoft Copilot, and DALL-E, has sparked considerable interest and adoption of AI. People have been applying generative AI to a wide range of contexts. Academic research is no exception. Within academia, AI has also drawn attention from researchers and educators (Dwivedi et al, 2023). Some of the key debates have centred around academic assessment integrity (Stokel-Walker, 2022; Eke, 2023). There are growing concerns around how ChatGPT and other generative AI applications could be used by students to produce assessments and consequently undermine academic integrity. While there is no one universally agreed solution and there is no validate tool to identify inappropriate use of generative AI , there is general consensus that a broader approach to integrating generative AI into pedagogy and assessment is required. It is important to note that while generative AI, such as ChatGPT, has been relatively new tool for research, other AI applications, asu cas Grammarly, has been utilised by a large number of users, indulging researchers, to improve professional and academic writing.

Due to the ability to process a large quantity of information and particularly processing natural language, generative AI has been experimented as a tool to research. Many researchers have utilised generative AI as part of the research methods in information systems research. Rossi et al. (2024) summarised that the current use of generative AI broadly fell into two categories, namely realistic experiments with generated content and using synthetic data. Many experiments require the use of text and images, and collecting naturally-occurred materials can be time-consuming and difficult, due to the need to be precise in controlling and measuring variables. Generative AI can help achieve realistic text and images to allow researchers to create materials efficiently. The other main use of generative AI for research is the creation of synthetic data. When it comes to the collection and storage of data related to human subjects, there are concerns about sensitive and identifiable information. Examples include synthetic profile pictures (Boyd at al., 2023) and synthetic medical images (Chambon et al., 2022). By using AI-generated synthetic data that closely resemble real-world data, researchers can mitigate privacy risks as well as address ethical and regulatory concerns.

Furthermore, generative AI has also been utilised to refine research questions and to check the completeness of the results (Burger et al., 2023). Going beyond text processing, AI has also been used as a tool for medical research (Castiglioni et al., 2021) such as recognising skin lesions with results matching or suppressing the accuracy of a dermatologist (Du‑Harpur et al., 2020). Literature review is another area where generative AI has also been tested as a tool (Aydın and Karaarslan, 2022;

Pan et al., 2023). Literature review is an essential part of research. A higher-level synthesis is necessary for a literature review, and it must incorporate ideas from other fields to provide a comprehensive overview of a given subject (Watson and Webster, 2020). It has been suggested that such tasks can be assisted by AI by generating an ontological map of concepts (Li et al, 2020). For instance, ChatGPT has been applied to automate the process of systematic literature review in the field of water and wastewater management (Alshami et al., 2023). ChatGPT has also been applied to generate research ideas in finance research, although it was considered that the literature synthesis and proposed testing frameworks could be further refined (Dowling & Lucey, 2023).

While generative AI could be a useful tool, it was also shown that when using ChatGPT as a tool to identify the literature for review, it could produce inaccurate and even non-existing results that could not be found in other databases (Haman & Školník, 2023). Additionally, although threats to academic integrity have always existed such as plagiarism, there are limited means currently for publishers to effectively identify the authenticity of authorship and the inappropriate use of generative AI as they would for plagiarism. Unquestionably generative AI applications will become more and more advanced and more readily available with major technology companies such as Micorsift and Google investing heavily in this area. The use of generative AI will consequently become more common in the coming months and years. Generally, it is considered that the use of generative AI as a tool in information systems research is still in a very early stage and more clear guidelines should be carefully developed (Rossi et al., 2024).

3.0 Approach

This study aims to explore the suitability of current generation AI applications commercially available for research purposes. The focus is on the generative AI's capability to synthesise information as a potential alternative or supplement to human-based information synthesisation. Thematic analysis for systematic literature review (Crossan and Apaydin, 2010; Tranfield et al., 2003) requires a large amount of information to be synthesised, and therefore it is selected as the domain of this exploratory experiment. ChatGPT4 is selected as the generative AI application, since it is one of the most widely used generative AI applications currently.

In order to evaluate the effectiveness of the thematic analysis produced by generative AI, this study compares the generative AI produced results with human-generated results (the benchmark), based on the same set of papers. The benchmark is peer-reviewed and therefore considered an appreciated baseline to compare against the generative AI produced result. This study uses the human-generated systematic literature review results from *Unlocking the value of artificial intelligence in human resource management through AI capability framework* (Chowdhury et al., 2023) published at Human Resource Management Review as the benchmark. This research paper conducted a systematic literature review and identified 18 themes based on 29 papers. This exploratory study will analyse 29 papers (from the Chowdhury et al., 2023 paper) following the same steps of identifying AI applications, barriers and drivers in HRM, using ChatGPT4 to identify and refine the key topics which will then be compared with the 18 topics in the baseline paper. The following section will describe the steps taken by this exploratory study and evaluate the generative AI produced results.

4.0 Steps and Results

The Chowdhury et al., 2023 paper used a systematic literature review protocol suggested in existing literature (Hopp et al., 2018; Tranfield et al., 2003). In their paper a topic modelling algorithm known as Latent Dirichlet Allocation was applied, resulting in 69 topics that were initially found from a Scopus search of relevant papers. After manual interpretation and text analysis were applied to the original 69 topics, 18 were then considered meaningful. Appendix B of the Chowdhury et al., 2023 paper provided the 18 topics and a sample of papers that demonstrated those topics.

4.1 Preparation

The first step in the process of analysing the sample literature review was to create a long summary in a narrative form (without headings) of each paper (500 words). This would allow the summaries to be combined for later analysis. It is feasible the summarising process was not needed as each paper in its entirety could be analysed in ChatGPT. However, when testing whole paper thematic analysis, the results were based on the paper headings and subheadings. For example, in Cubrics (2020) paper, '*Drivers, Barriers and Social Considerations for AI Adoption in Business and*

Management: A Tertiary Study", the following "*topics*" were identified through ChatGPT:

- Introduction to the Study
- Rise of AI
- Research Gap
- Methodology
- Research Questions
- Importance of Social Factors
- Conclusion and References

Therefore, it is considered not suitable to use this prompt to identify topics, as ChatGPT focuses on the heading in the paper as topics. Consequently we used ChatGPT to summarise the paper to generate a narrative for each paper for next steps of analysis. The prompt "*summarise the paper into 500 words*" was used, as it generated an appropriate overview of the paper. The summaries were then used for the rest of the experiment.

4.2 Applications of AI in HRM

To repeat the analysis carried out by Chowdhury et al (2023) to understand the applications of AI in HMR, we imported the summary of all 29 sample papers into ChatGPT and provided the following prompt: "*provide a list of applications of AI in HRM reported in this literature*". This resulted in the following results (Table 1, column 2), which are shown against the results from Chowdbury et al (2023) in Table 1, column 1.

Original list of applications of AI in HRM	**Generative AI analysis of literature**
Candidate Experience (Job Applications)	Talent Acquisition
Candidate Recruitment	Employee Engagement Analysis
Onboarding	Performance Evaluation
Employee Engagement	Predictive Modelling
Career Development	Training and Development
Employee Performance Appraisal	Enhancing Decision Making
Compensation Packages	Cultural Shift Towards AI in SBMS
Employee Skills Development	Ethical Considerations in HR Practices

Workforce Management Analytics	Support For Administrative HR Tasks
HR Budget and Resource Allocation	Workforce Analytics

Table 1. List of applications of AI in HRM

4.3 Drivers of AI in HRM

The next step in our research was to look for the drivers to AI adoption. In the benchmark paper the identification of drivers was not presented as an example table or list. Consequently the list outlined in table 2 has been created through human interpretation of section 3.4 in the original paper. To conduct the same analysis with our sample data, the summaries of our sample articles were inputted into ChatGPT with the prompt: "provide a list of drivers of AI adoption in HRM reported in this literature". This is shown in Table 2, column 2 against the benchmark list in column 1.

Original list of drivers of AI in HRM	**Generative AI analysis of literature**
Objectivity and Accuracy	Technological Advancements
Enhancing Creativity and Innovation	Demand for Data-Driven Decision-Making
Streamlining Organisational Processes	Enhanced Employee Experience and Engagement
Recruitment and Talent Acquisition	Efficiency and Automation of Administrative Tasks
Reducing Bias in Candidate Screening	Talent Acquisition and Management
Monitoring and Performance Measurement	Predictive Analytics in Workforce Planning
Predicting HR Development	Personalisation of Learning and Development
Employee Retention	Employee Performance Evaluation
Operational Optimisation	Cultural Transformation and Change Management
Enhancing Dynamic Capabilities	Global and Cross-Border HR Management
Employee Well-being	Compliance and Regulatory Requirements
	Strategic HRM and Decision Support

Table 2. List of drivers of AI in HRM

4.4 Barriers of AI in HRM

Finally, we looked at the barrier to AI adoption in HRM. The analysis on this topic occurred in section 3.5 of the benchmark paper and again this requires human interpretation of the findings to produce a list of barriers. This is shown in table 3 column 1. To conduct the same analysis with our sample data, the combined summaries of our sample of articles was inputted into ChatGPT with the prompt: "*provide a list of barriers to AI adoption in HRM reported in this literature*". This is shown in Table 3 column 2.

Original list of barriers of AI in HRM	**Generative AI analysis of literature**
Complexity of HR Phenomena	Data Privacy Concerns
Small Data	Talent Gap in Analytics Skills
Ethical Constraints	Cultural Resistance to Change
Employee Reactions	Ethical Implications and Bias
Privacy and Data Protection	Integration with Existing Systems
Constant Tracking Issues	Limited Access to Technology
Potential Bias in Algorithms	Digital Divide
Data Quality Assessment	Cost and Resource Allocation
Training Dataset Optimisation	Lack of Clear Business Case
Technological Integration	Regulatory and Compliance Issues
Developing a Data-Centric Culture	Uncertainty about AI Capabilities and Outcomes
Technology Turbulence	Misalignment between AI Solutions and Organisational Needs
Transparency and Interpretability	Employee Privacy and Consent Concerns
Epistemological Issues in AI-Driven Recruitment	Need for Interdisciplinary Collaboration
AI's Limitations in Creative and Social Intelligence	Job Security Fears among Employees
External Environmental Variables	
Human-AI Synergy Requirement	

Table 3. List of barriers of AI in HRM

5.0 Findings and Conclusion

Based on the experiment, it is considered that the summaries generated by ChatGPT were suitable, as they tended to capture the essence of the papers while providing more information than the abstracts of the papers. This allowed the experiment to continue to the next step of using generative AI to identify topics prior to comparing generative AI and human results. When it comes to identifying topics based on the paper summaries, generative AI produces some results that are highly similar to the human-generated results. For example, 'Recruitment and Talent Acquisition' was one of the topics identified by humans and it can be closely matched to 'Talent Acquisition and Management' which was produced by ChatPGT. Similarly, 'Workforce Management Analytics' can be closely matched to 'Workforce Analytics', and 'Employee Engagement' to 'Employee Engagement Analysis'. This high level similarly can be observed for all three sets of experiments. This is not a surprise, as generative AI has been found to produce content that can be indistinguishable from human-generated content.

On the other hand, there are also some mismatches between human and generative AI results. For instance,'Predictive Modelling' was one of the topics as AI applications in HRM identified by ChatPGT that cannot be linked to topics identified by humans. We considered that the term 'Predictive Modelling' is too broad as it could overlap with 'Workforce Analytics', which was also identified by ChatGPT. It is likely that humans would synthesise elements of predictive modelling into the context where predictive modelling was applied to, e.g. sales prediction or workforce planning.

Additionally, another interesting finding from this study was that similar numbers of topics were identified, without guiding prompts of the expected number of topics. For Applications of AI in HRM, both humans and generative AI produced 10 topics. For drivers of AI in HRM, humans identified 11 topics and generative AI produced 12 topics. For the barriers of AI in HRM, humans identified 17 topics and generative AI identified 15 topics. Based on this exploratory study, generative AI was able to synthesise the papers and narrow down the topics into a very similar number of items to the human results.

It is also noted that the prompt word 'topics' did not generate suitable results when the study tried to use ChatGPT to identify the key topics of each paper. ChatGPT picked up the headings, e.g. methodology, as the topics. While ChatGPT was able to identify topics from a smaller text body, i.e., 500-word summary of a paper, it did not identify appropriate topics from a larger text body, e.g. a full research paper. It is possible that different prompt words or sets of prompt words/questions could lead to more effective results when analysing a larger text body. This should be further explored and evaluated.

The results presented here are the initial findings in the exploratory study. As an exploratory study of a relatively small scale, the findings cannot be over generalised. Future work would consider the inter-rater reliability between the human reviewers and ChatGPT on a more significant scale, with a more complex set of prompt questions. This comparison would be key to understanding the difference between human and generative AI in its categorisation of papers into topics, as well as how scholars could utilise generative AI to synthesise a large amount of literature, which could potentially accelerate the speed of systematic literature review.

References

Alshami, A., Elsayed, M., Ali, E., Eltoukhy, A. E., & Zayed, T. (2023). Harnessing the Power of ChatGPT for Automating Systematic Review Process: Methodology, Case Study, Limitations, and Future Directions. *Systems*, 11(7), 351.

Aydın, Ö., & Karaarslan, E. (2022). OpenAI ChatGPT generated literature review: Digital twin in healthcare. Available at SSRN 4308687.

Boyd, A., Tinsley, P., Bowyer, K., & Czajka, A. (2023, June). The value of ai guidance in human examination of synthetically-generated faces. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 37, No. 5, pp. 5930-5938).

Burger, B., Kanbach, D. K., & Kraus, S. (2023). The role of narcissism in entrepreneurial activity: a systematic literature review. *Journal of Enterprising Communities: People and Places in the Global Economy*, (ahead-of-print).

Castiglioni, I., Rundo, L., Codari, M., Di Leo, G., Salvatore, C., Interlenghi, M., ... & Sardanelli, F. (2021). AI applications to medical images: From machine learning to deep learning. *Physica Medica*, 83, 9-24.

Chambon, P., Bluethgen, C., Langlotz, C. P., & Chaudhari, A. (2022). Adapting pretrained vision-language foundational models to medical imaging domains. arXiv preprint arXiv:2210.04133.

Chowdhury, S., Dey, P., Joel-Edgar, S., Bhattacharya, S., Rodriguez-Espindola, O., Abadie, A., & Truong, L. (2023). Unlocking the value of artificial intelligence in human resource management through AI capability framework. *Human Resource Management Review*, 33(1), 100899.

Crossan, M.M. and Apaydin, M., 2010. A multidimensional framework of organizational innovation: A systematic review of the literature. *Journal of Management Studies*, 47(6), pp.1154-1191.

De Cremer, D., & Kasparov, G. (2021). AI should augment human intelligence, not replace it. *Harvard Business Review*, 18, 1.

Du‐Harpur, X., Watt, F. M., Luscombe, N. M., & Lynch, M. D. (2020). What is AI? Applications of artificial intelligence to dermatology. *British Journal of Dermatology*, 183(3), 423-430.

Dowling, M., & Lucey, B. (2023). ChatGPT for (finance) research: The Bananarama conjecture. *Finance Research Letters*, 53, 103662.

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... & Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, 102642.

Eke, D. O. (2023). ChatGPT and the rise of generative AI: Threat to academic integrity?. *Journal of Responsible Technology*, 13, 100060.

Haman, M., & Školník, M. (2023). Using ChatGPT to conduct a literature review. *Accountability in Research*, 1-3.

Hopp, C., Antons, D., Kaminski, J., & Oliver Salge, T. (2018). Disruptive innovation: Conceptual foundations, empirical evidence, and research opportunities in the digital age. *Journal of Product Innovation Management*, 35(3), 446-457.

Li, J., Larsen, K., & Abbasi, A. (2020). TheoryOn: A design framework and system for unlocking behavioral knowledge through ontology learning. *MIS Quarterly*, 44(4).

Liu, R., Gupta, S., & Patel, P. (2023). The application of the principles of responsible AI on social media marketing for digital health. *Information Systems Frontiers*, 25(6), 2275-2299.

Mahdi, S. S., Battineni, G., Khawaja, M., Allana, R., Siddiqui, M. K., & Agha, D. (2023). How does artificial intelligence impact digital healthcare initiatives? A review of AI applications in dental healthcare. *International Journal of Information Management Data Insights*, 3(1), 100144.

Manis, K. T., & Madhavaram, S. (2023). AI-Enabled marketing capabilities and the hierarchy of capabilities: Conceptualization, proposition development, and research avenues. *Journal of Business Research*, 157, 113485.

Makarius, E. E., Mukherjee, D., Fox, J. D., & Fox, A. K. (2020). Rising with the machines: A sociotechnical framework for bringing artificial intelligence into the organization. *Journal of Business Research*, 120, 262-273.

Metz, A. (2023). exciting ways to use ChatGPT–from coding to poetry. TechRadar.

Pan, S. L., Nishant, R., Tuunanen, T., & Nah, F. F. H. (2023). Literature review in the generative AI era-how to make a compelling contribution. *The Journal of Strategic Information Systems*, 32(3).

Rahman, M., Ming, T. H., Baigh, T. A., & Sarker, M. (2023). Adoption of artificial intelligence in banking services: an empirical analysis. *International Journal of Emerging Markets*, 18(10), 4270-4300.

Rossi, S., Kwon, Y., Auglend, O.H., Mukkamala, R.R., Rossi, M., & Thatcher, J. (2023). Are Deep Learning-Generated Social Media Profiles Indistinguishable from Real Profiles? *Proceedings of the 56th Hawaii International Conference on System Sciences*, 134–143. https://hdl.handle.net/10125/102645

Rossi, S., Rossi, M., Mukkamala, R. R., Thatcher, J. B., & Dwivedi, Y. K. (2024). Augmenting research methods with foundation models and generative AI. *International Journal of Information Management,* 102749.

Stokel-Walker, C. (2022). AI bot ChatGPT writes smart essays-should academics worry?. *Nature*.

Stokel-Walker, C. (2023). ChatGPT listed as author on research papers: many scientists disapprove. *Nature*, 613(7945), 620-621.

Thieme, A., Hanratty, M., Lyons, M., Palacios, J., Marques, R. F., Morrison, C., & Doherty, G. (2023). Designing human-centered AI for mental health: Developing clinically relevant applications for online CBT treatment. *ACM Transactions on Computer-Human Interaction*, 30(2), 1-50.

Tranfield, D., Denyer, D. and Smart, P., 2003. Towards a methodology for developing evidence‐informed management knowledge by means of systematic review. *British journal of management*, 14(3), pp.207-222.

van Dis, E. A., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. (2023). ChatGPT: five priorities for research. *Nature*, 614(7947), 224-226.

Watson, R. T., & Webster, J. (2020). Analysing the past to prepare for the future: Writing a literature review a roadmap for release 2.0. *Journal of Decision Systems*, 29(3), 129-147.

Yüksel, N., Börklü, H. R., Sezer, H. K., & Canyurt, O. E. (2023). Review of artificial intelligence applications in engineering design perspective. *Engineering Applications of Artificial Intelligence*, 118, 105697.

Zhuo, T.Y., Huang, Y., Chen, C. and Xing, Z., 2023. Exploring ai ethics of chatgpt: A diagnostic analysis. arXiv preprint arXiv:2301.12867.