Human Interaction and Emerging Technologies (IHIET-AI), Vol. 201, 2026, 332–338
https://doi.org/10.54941/ahfe1007187

Comparison of Human and AI-Driven Interview Data Analysis in Industrial Work Context

Antti Tammela, Susanna Aromaa, and Hanna Lammi

VTT Technical Research Centre of Finland Ltd., Visiokatu 4, Tampere FI-33720, Finland

ABSTRACT

The use of artificial intelligence (AI) in research has increased. This study compares human and AI approaches to analyzing qualitative field trial interview data in human-centered design. The results show that both the human analysis and large language models (ChatGPT 5 Pro) analysis find largely the same key findings. However, some of the findings differed, mostly at the level of perspective and abstraction. For example, AI emphasized safety and security issues more than human analysis. The data also revealed a clear AI interpretation error related to linking a participant's comment to the wrong technology. This highlights the importance of both expert validation and careful prompt design. In conclusion, the study suggests that human and AI analysis do not replace but complement each other. The most promising solutions are found in a hybrid model, where the speed and systematicity of AI are combined with the human analyst ethical judgment and knowledge on the interview data.

Keywords: Artificial intelligence, Interview analysis, Human-centred design, Industrial field trial, Generative AI, Compare

INTRODUCTION

The rapid development of artificial intelligence (AI), especially in the form of large language models (LLMs), has opened new possibilities for the analysis of qualitative data. This has traditionally relied on the expertise, contextual understanding and interpretation of human analysts. There is growing interest in the field of human-centred design on whether analytical processes can be accelerated and systematized with the help of AI. This must be done without compromising the quality and reliability of the results. Empirical evidence on the differences between human and AI-based analyses in real industrial environments is still limited. This article compares the analyses of interview data from an industrial work context by human analysts and AI-based methods. During the trial, test-users conducted the turning of a gear wheel with the assistance of an AI and a remote expert. The test-users were interviewed afterwards. After the trial, the AI was guided by prompts to carry out the analysis according to the same principles as the human researcher performed it. The goal was to identify the key similarities and differences between human and AI data analyses in industrial work context.

Received February 5, 2026; Revised February 26, 2026; Accepted March 5, 2026; Available online April 1, 2026

Artificial Intelligence in Research Work

Artificial intelligence is used increasingly in research work. AI offers tools that are naturally suited to research and make it more efficient. However, the reliability of AI has been highlighted in several studies (Qiao et al., 2025; Han et al., 2025). Can we trust that it will not hallucinate at any point? Can AI interpret research results as well as humans? Despite this, LLM-based tools are increasingly used for coding, identifying themes, and summarizing results across disciplines. At the same time, both academic research and industry case studies emphasize the need to maintain the role of the human researcher as the final authority for interpretation, reliability, and ethical evaluation (Costa et al., 2025; Blanchard et al., 2025).

AI and LLM's are quickly reshaping how researchers conduct surveys and experiments (Aromaa et al., 2024; Takaffoli et al., 2024; Tholander et al., 2023), AI is reviewing literature and designing instruments for administering studies. It is coding data, and interpreting results. These tools offer novel opportunities to improve research productivity and advance methodology. However, with this potential comes a critical challenge: researchers often use these systems without fully understanding how they work. Transparency (prompts, templates, and settings) and repeatability are emphasized. Participants' privacy and copyright must be respected. Results should be systematically validated when AI is used for measurement and open-text coding.

AI in Interview Data Analyses

Danis et al. (2024) found that AI (GPT-4) can identify key themes in qualitative analysis of patient interviews, such as urinary, sexual, mental health, and hygiene problems, quite well. The agreement with human analysts was quite good. But humans consistently find more subtle and detailed subthemes. AI analysis is fast and relatively reliable at the level of general entities. Detection of rarer or more nuanced issues varies from analysis to analysis. The conclusion of the study is that AI serves as a useful and time-saving tool in qualitative research but does not replace human analysis. The best results are achieved by combining the strengths of both (Danis et al., 2024; Refosco et al., 2025).

Prescot et al. (2024) compared the ability of humans and AI (ChatGPT and Bard/Gemini) to perform qualitative thematic analysis of short text messages. The researchers used both inductive and deductive analysis. They compared how well AI and humans found the themes and coded the text in the same way. All AI models were able to find most of the same themes as humans. But detailed coding was only moderate or reliable. Humans were better at identifying subtle and interpretive meanings. The biggest difference was in efficiency: AI took an average of only about 20 minutes to analyze, compared to nearly 10 hours for humans. The authors suggest that the best solution is a hybrid model. AI speeds up the work, but humans ensure the accuracy and ethics of the analysis (Prescot et al., 2024).

Mathis et al. (2024) compared the inductive thematic analysis produced by a locally run open LLM with human coding. This was carried out on data from 21 interviews conducted in a psychiatric setting. In the results, LLM found 6 themes from clients and 5 from clinicians. In the human analysis, there were 6 and 7 corresponding themes. The similarity of themes between human and LLM ranged from moderate to significant. This suggests that open LLMs can effectively support qualitative research. However, they emphasize the need for human control (Mathis et al., 2024).

METHODS

The aim of the study was to assess the suitability of AI for the analysis of interview data. The interview material had been collected in a study examining the use of emerging technologies in industrial work. The objective was to determine whether AI analysis produces results comparable to those obtained through human analysis. The study paid great attention to ensuring that the analysis performed by both AI and humans was done in the same way.

Data to be analyzed was collected in a user study, in which five workers performed a lifting task using a crane. They used two novel technologies: (1) real-time guidance provided by a remote expert via see-what-I-see smart glasses (Iristick) and (2) guidance provided by an AI assistant. Finally, a semi-structured interview explored both technologies in depth, covering their strengths, challenges, future potential in industrial work, and their influence on knowledge sharing, information flow, and the meaningfulness of work, while also addressing any concerns raised by participants. The analysis of the interviews was first done by two human factors researchers with over 20 years of experience in user studies. The same researchers had previously interviewed the employees. They analyzed the results deductively based on a semi-structured questionnaire framework which focused on themes such as (1) positive feedback, (2) challenges and development ideas, (3) meaningful work (beneficence, relatedness, competence, identity, autonomy, embodiment), and (4) knowledge sharing (Fereday et al., 2006; Elo and Kyngäs, 2008). The same division was used for both the AI assistant and the remote assistant results.

AI analysis (AIA) was performed according to the same specifications as the human analysis (HA). As for background information, AI was given a description of the two technologies being tested. An AI prompt was implemented to provide the results of the analysis with the same main topics and format as the HA analysis. The prompts asked the AI to provide the results in a similar slide format as the human analysis. This helped with the qualitative comparison of results afterwards. AI was prompted to perform the analysis as if it was a well-experienced and trained human factors researcher. ChatGPT 5 pro was used for the analysis. The AI was provided with transcripts of the interviews.

The results of the analyses were compared using qualitative methods. Both results were first collected into slide sets. From these, the results were transferred to a table, where both analysis results were next to each other for each 4 main result area. After this, a researcher who developed the AI system evaluated the similarity of each theme. The results were analyzed by comparing four main themes from both technologies (AI assistant, Smart glasses) qualitatively. Finally, these were compiled into overall results.

RESULTS AND DISCUSSION

Both AIA and HA analyses produced very similar results (see Table 1). Both highlighted, for example, the impact of factory noise. This finding is consistent with Mathis et al (2024)'s study. The differences in methods were primarily related to perspective and level of abstraction, not to opposing outcomes. The results showed that there was at least one difference in each main topic. The analyses were most similar in terms of usability and problem detection. The biggest differences were related to meaningfulness of work and knowledge, for example, HA emphasized the importance of tacit knowledge transfer through AI assistants. However, the differences were largely interpretative. Human analysis focused on barriers to workflow and practical problems, while AIA translated the same problems into design models and error correction steps. Direct quotes were similar in both methods. They brought up the same result. The advantage of AI was that it provided a direct link to the transcribed interview with a timestamp. This made it easier to check for accuracy.

Table 1: Human and AI comparison examples.

Topic	Examples of Similar Findings	Examples of Findings Obtained Solely Through One of the Analysis
Positive feedback	Guiding, teaching novice workers, and onboarding new employees Typing interaction felt accurate and natural	Tasks can be recorded for later use (HA). Sometimes AI assistant can be clearer than outsourced IT support (AIA).
Challenges and development ideas	Factory noise masked speech	Add safety warnings (load can sway/ rotate) (AIA).
Meaningful work	Collaboration reduced 'working alone' feeling	AI introduces new dimensions and value to work (HA). Potential to lower stress in unfamiliar tasks (AIA). Data-privacy reduce willingness to ask freely (AIA).
Knowledge sharing	Consistent information sharing	Tacit knowledge sharing (HA) Good for remote audits or sign-off without travel (AIA).

Although both methods produced similar main results, there were differences in details. Often AI highlighted more safety and security issues, For example, AI highlighted the need for safety warnings that emerged in the interviews (Table 1). This suggests that AI can highlight themes that are less prominent in human analysis. Even if they do not fundamentally change the overall direction of the conclusions. The differences arise also in the way analysis express the results. Human analysis lists problems: AI translates them into design models and error correction steps. At the same time, it shows that different analysis approaches can reinforce each other.

The comparison of analyses also revealed one clear misinterpretation in AIA. One response was incorrectly attributed to the use of smart glasses, even though the comment was related to AI assistant. This issue could also have occurred with human analysts if they had relied solely on the transcripts. Since the two HF researchers conducted the interviews themselves, they had a better understanding of aspects that are not always explicitly stated. In other words, transcripts may not capture all the tacit knowledge conveyed during the interview. This highlights two essential conclusions. First, AI analysis may not be sufficient on its own, and the results it produces must be checked by an expert, preferably someone who has participated in the interviews. Second, the quality of the analysis strongly depends on the instructions (prompts) provided. Unclear or incomplete descriptions of technologies can lead to incorrect conclusions.

Together, the results suggest that human and AI analysis are not substitutes for each other, but complementary approaches as Danis et al. (2024) also note. AI can speed up analysis and support the structuring of themes. Human analysis, in turn, ensures that interpretations are contextually reliable, human-centered, and ethically justified (Danis et al., 2024; Mathis et al., 2024; Prescot et al., 2024).

This study has certain limitations to consider. The data set was small, as only five individuals were interviewed. Additionally, both the AI tool development and the comparative analysis were conducted by a single researcher. However, this was a pilot study, and it would be beneficial in the future to replicate the research with a larger participant group and to employ quantitative evaluation methods for comparing the results.

CONCLUSION

Interest is increasing in how AI might speed up and organize analytical processes, particularly in human-cantered design research. This study compared the results of human analysis and AI analysis on interview data from industrial workers who performed a work task by utilizing emerging technologies such as AI and smart glasses.

The results show that the analysis methods are highly consistent in terms of the found features. Both identified the same key themes, such as strengthening the sense of teamwork and novice orientation. The key conclusion of this study is that a hybrid analysis model is promising. Human interpretation and fast, structured AI analysis could be combined in the research process. At the same time, it is necessary to recognize the limits of AI. The results should be checked afterwards, and the analysis prompts must be done carefully. In this way, misleading results and incorrect information about the technology being studied can be avoided. The findings show that AI can be used to enhance and diversify analysis. The results of the study can also be utilized in an industrial context.

ACKNOWLEDGMENT

This project was supported by the Business Finland (5785/31/2023). The authors are grateful to all researchers and project partners who have contributed to and supported the work presented in this publication.

REFERENCES

Aromaa, S., Lammi, H. 2024. Exploring the Potential of Generative Artificial Intelligence in Human-Centred Design. In: Tareq Ahram, Waldemar Karwowski, Darko Etinger and Tea Mijač (eds) Human Systems Engineering and Design (IHSED2024): Future Trends and Applications. AHFE (2024) International Conference. AHFE Open Access, vol-2. AHFE International, USA. http://doi.org/10.54941/ahfe1005523

Blanchard, S., Duani,N., Garvey, A., Netzer, O. and Travis Tae. 2025. New Tools, New Rules: A Practical Guide to Effective and Responsible Generative AI Use for Surveys and Experiments in Research. 2025. Journal of Marketing, Nov2025. Business Source Elite. https://doi.org/10.1177/00222429251349882

Costa, A., Bryda, G., ChristouP., and Kasperiuniene J. 2025. AI as a Co-researcher in the Qualitative Research Workflow: Transforming Human-AI Collaboration. International Journal of Qualitative Methods,10/14/2025. (Complementary Index). https://doi.org/10.1177/16094069251383739

Danis, l., Fernandez, A., Schwartz, R., Rios, N., Carlisle, M., N., Amend, G., A., Patel, H., V. and Breyer, B., N. 2024. Comparing GPT-4 and Human Researchers in Health Care Data Analysis: Qualitative Description Study. Journal of Medical Internet Research. Supplemental Index. doi: 10.2196/56500

Elo, S., & Kyngäs, H. 2008. *The qualitative content analysis process. Journal of Advanced Nursing, 62*(1), 107–115.* DOI: 10.1111/j.1365-2648.2007.04569.x

Fereday, J., and Muir-Cochrane, E. 2006. Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. 2006. International Journal of Qualitative Methods, 5(1), 80–92. https://doi.org/10.1177/160940690600500107

Han, J and Yue, Z. 2025. Integrating AI Into Methodology: A Critical Review of a Research Methods Course in Postgraduate Interdiciplinary Built Environment Education. 2025 5th International Conference on Artificial Intelligence and Education (ICAIE), Artificial Intelligence and Education (ICAIE), 2025 5th International Conference on, 20250514. IEEE Xplore Digital Library.

Mathis, W., Zhao, S., Pratt, N., Weleff, J. and De Paoli, Stefano. 2024. Inductive thematic analysis of healthcare qualitative interviews using open-source large language models: How does it compare to traditional methods? 2024. Computer Methods and Programs in Biomedicine. ScienceDirect.

Prescott, M., Yeager, S., Ham, L., Saldana, C, Serrano, V., Narez, J. Paltin, D., Delgado, J. Moore, D. and Montoya, J. 2024. Comparing the Efficacy and Efficiency of Human and Generative AI: Qualitative Thematic Analyses. 2024. JMIR AI; 2024, Vol. 3, p1–13, 13p. Complementary Index.

Qiao, S., Fang, X, Wang, J., Zhang, R., Li, X. and Kang, Y. 2025. Generative AI for thematic analysis in a maternal health study: coding semistructured interviews using large language models. 16 May 2025, https://doi.org/10.1111/aphw.70038.

Refosco, P. and Crowther, K. Automating Voice of Customer Analysis with AI Workflows Built on GPT-4o-Mini Model. 2025. IEEE International Conference on AI and Data Analytics (ICAD), AI and Data Analytics (ICAD), 2025 IEEE International Conference on, 20250624. IEEE Xplore Digital Library.

Takaffoli, M., Li, S., & Mäkelä, V. 2024. Generative AI in User Experience Design and Research: How Do UX Practitioners, Teams, and Companies Use GenAI in Industry? Proceedings of the 2024 ACM Designing Interactive Systems Conference,

Tholander, J., & Jonsson, M. (2023, July. Design ideation with AI-sketching, thinking and talking with generative machine learning models. In Proceedings of the 2023 ACM designing interactive systems conference (pp. 1930–1940).