

# NLP AND GENAI ANALYSIS: AUTOMATING THEMATIC ANALYSIS AND CODING OF QUALITATIVE EDUCATIONAL RESEARCH DATA

James D. Cleaver<sup>a</sup>, Laura K. McKemmish<sup>a</sup>, Sara H. Kyne<sup>a</sup>

Presenting Author: James D. Cleaver ([j.cleaver@student.unsw.edu.au](mailto:j.cleaver@student.unsw.edu.au))

<sup>a</sup>School of Chemistry, Faculty of Science, University of New South Wales, Sydney NSW 2052, Australia

**KEYWORDS:** Natural Language Processing, Generative AI, Qualitative Analysis, Automated Coding, Inter-Rater Reliability

## BACKGROUND

Qualitative analysis of large datasets often involves manual coding and theme extraction, a process that is both time-consuming and subject to human interpretation. Traditional approaches to ensuring inter-rater reliability, where multiple analysts independently manually code data to achieve consensus, can be particularly laborious. Advances in Natural Language Processing (NLP) and Generative AI (GenAI) offer promising alternatives to streamline these processes by automating theme extraction and coding (Fitkov-Norris & Kocheva, 2023). Improving the efficiency and accuracy of qualitative analysis is crucial in Science Education Research, where effectively interpreting student open-ended survey responses and interviews is essential for promoting engagement with scientific concepts and research tasks (Sinatra et al., 2015).

## AIMS

This study aims to explore the efficacy of NLP and GenAI techniques in expediting qualitative analysis. By automating the identification of themes and the application of coding frameworks, we aim to reduce the need for complete manual coding and inter-rater reliability checks, thus making qualitative analysis more efficient and scalable.

## DESIGN AND METHODS

A suite of NLP and GenAI tools will be used to analyse qualitative data from open-ended survey responses and interviews. The process involves configuring and engineering computer models to recognise and extract themes and apply standardised coding frameworks to categorise data accurately. These models' performance will be evaluated against independently manually coded data to assess their accuracy and reliability. Additionally, we will examine the models' ability to handle different data types, ensuring robustness across different qualitative research scenarios.

## RESULTS AND CONCLUSIONS

The study will trial and evaluate a range of AI techniques to investigate the potential of NLP and GenAI in reducing the time and manual coding required for qualitative analysis while maintaining high levels of accuracy. The automated application of coding frameworks will be assessed for its ability to replicate the consistency of current manual coding techniques, as measured through inter-rater reliability checks, thereby potentially removing the need for multiple researcher inter-rate reliability checks in the future. The intent of this research is to inform and enhance qualitative research workflows, allowing researchers to focus more on interpreting and leveraging their findings to understand and improve student motivation and experience.

## REFERENCES

- Fitkov-Norris, E., & Kocheva, N. (2023). Are we There yet? Thematic Analysis, NLP, and Machine Learning for Research. *European Conference on Research Methodology for Business and Management Studies*, 22(1), Article 1. <https://doi.org/10.34190/ecrm.22.1.1616>
- Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The Challenges of Defining and Measuring Student Engagement in *Science. Educational Psychologist*, 50(1), 1–13. <https://doi.org/10.1080/00461520.2014.1002924>

Proceedings of the Australian Conference on Science and Mathematics Education, The University of Canberra, 18 – 20 September 2024, page 22, ISSN 2653-0481.