TRACE: Transforming Qualitative Public Health Analysis
AI-Powered Thematic & Sentiment Extraction for Faster, Deeper Insights

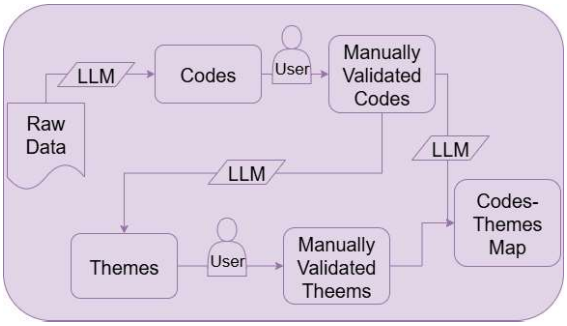
Isobel Smith
UKHSA, Health Analytics & Automation Team

Introduction

TRACE (Thematic Recognition And Contextual Extraction) is an AI tool built on the secure Llama 3.3-70B language model. It speeds up the analysis of free-text data from surveys, interviews, and questionnaires, turning qualitative feedback into actionable intelligence for public health decision-making. TRACE is designed to address real-world challenges, delivering rapid insights where they are needed most. TRACE assists analysts by rapidly surfacing themes and sentiment, enabling faster interpretation.

Traditional qualitative analysis creates significant barriers:
- Manual coding is time-intensive and costly
- Limited accessibility
- Delayed insights during health emergencies
- Scalability challenges with large datasets

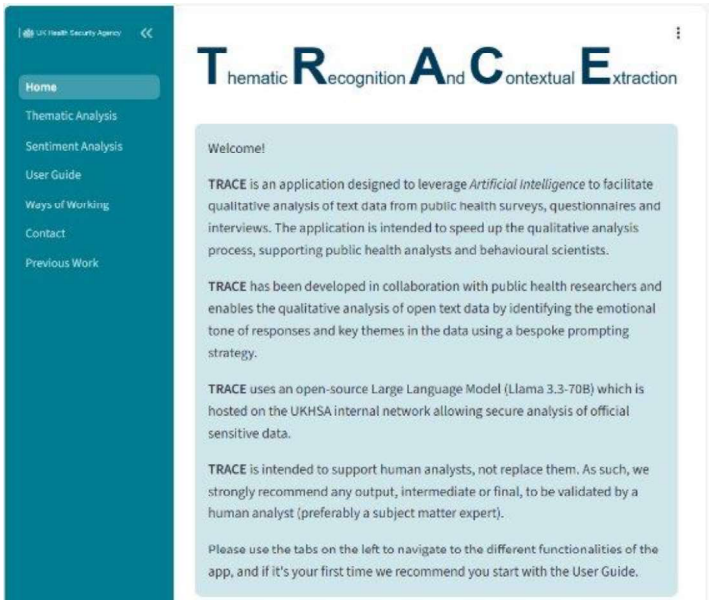
TRACE helps to democratise qualitative research, making it faster, more accessible, and more impactful.



Methodology

How TRACE Works: From Data to Action
Input: Upload anonymised text data (e.g., survey responses, interview transcripts).
Process: TRACE's secure AI engine performs two key types of analysis:
- Thematic Analysis: Identifies and tracks frequency of key themes and emerging issues.
- Sentiment Analysis: Detects emotional tone (Positive, Negative, Neutral) within responses.

Output: Clear, interpretable visualisations.



Results

TRACE is a significantly impactful time saving tool:
- Time taken reduced by 2/3 compared to human analyst.
- Significant resource saving as creation takes most mental effort.

Use Case 1: Positive Voices Survey (HIV)
Challenge: Manually analysing thousands of complex patient experience surveys to understand barriers to care, stigma, and quality of life.
TRACE Action: Expedited thematic analysis identified key emergent themes and map sentiment to show how experiences vary by demographic.
Performed on 1602 survey response with a coding step which used a scoring rubric developed with the subject matter experts.
Impact: 10% of codes were rejected (over-coding).
3% of responses had missing codes found by the human (under-coding).
Enable rapid, data-driven adjustments to support services and policy, directly improving care for people living with HIV.

Use Case 2: COVID-19 Pandemic Learning
Challenge: During the pandemic, health systems were inundated with qualitative data that was too vast to analyse manually in time to inform response.
TRACE Action: Process large volumes of text data, such as internal surveys and semi-structured interviews to identify emerging concerns.
10 out of 11 manual themes (91%) and 23 of 43 manual subthemes (53%) were directly mapped to by the LLM themes and subthemes compared with fully manual thematic analysis.
Impact: Provide decision-makers with agile insights into public perception and operational bottlenecks, strengthening preparedness for future health emergencies.

Conclusion and Future Directions

TRACE moves beyond theory to proven application. By learning from the depth of patient experience, TRACE is an essential tool for building a more responsive, evidence-based, and equitable public health system. This is because faster qualitative insight enables quicker, more impactful public health action.

Key Benefits for Public Health Teams:
Speed: Reduce analysis time from months to days, accelerating intervention and response.
Scale: Analyse datasets of any size consistently, without manual fatigue.
Accessibility: No coding required. Empowers behavioural scientists, policy teams, and analysts.
Trust & Security: Developed with government-backed security protocols for sensitive patient data

TRACE's key accomplishments are:
- Potentially Transformative Tool: TRACE can change how qualitative health data is analysed
- Evidence-Based Policy: Enables faster, more responsive public health interventions
- Democratic Access: Can help to empower more non-technical researchers to conduct sophisticated analysis
- Scalable Solutions: Supports analysis from small studies to population-level datasets

Future Development: Expansion to additional healthcare domains, such as health protection policy development and knowledge and library services enhancement.

Acknowledgements and References

Acknowledgements: UKHSA (especially Tom White and the Health Analytics and Automation Team), HDR UK Internship Programme.
UKHSA Guidelines, Meta Llama 3.3, Positive Voices Survey.
Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa
Minaee, Shervin, Tomas Mikolov, Narjes Nikzad, Meysam Chenaghlu, Richard Socher, Xavier Amatriain, and Jianfeng Gao. "Large language models: A survey." arXiv preprint arXiv:2402.06196 (2024).