*CERN IdeaSquare Journal of Experimental Innovation, 2026; 10(2):141-143*
*DOI: https://doi.org/10.23726/cij.2026.1827*

***METHODOLOGICAL NOTE***

# Human Interpretation, AI Support: A Workflow for Thematic Analysis using Large Language Models in Co-design

Crystal A. Silver,[1] Diane Morrow[1], Stefano De Paoli[1*]

[1] 1st Department of Sociological and Psychological Sciences, Abertay University, Dundee, United Kingdom
*Corresponding author: s.depaoli@abertay.ac.uk

## WHAT IS IT ABOUT?

This methodological note describes a workflow for thematic analysis in which Large Language Models (LLMs) are used to support specific analytic tasks on Co-creation and Co-design user data. LLMs are advanced artificial intelligence systems trained on large volumes of textual data that can generate articulated textual responses when prompted. The primary benefit of this workflow is the streamlining of time-intensive analytic tasks, enabling Co-creation and Co-design teams to maintain methodological rigour in qualitative analysis while working under tight time constraints.

In the qualitative research literature, Co-creation is commonly defined as the collaborative generation of knowledge by academics and non-academic stakeholders working together across stages of a project, including the interpretation and application of findings (Pearce et al., 2021). From this perspective, the workflow supports Co-creation by structuring how analytic artefacts are produced, discussed, and validated across a research team and wider stakeholder group.

The workflow was developed in the context of qualitative user research and Co-design conducted to inform the design of interdisciplinary research discovery platforms (Morrow et al., 2025). Semi-structured interview data with potential end-users was collected and then analysed using thematic analysis, as an initial step of Co-design to capture the user-needs. LLMs were introduced at selected stages to assist with pattern processing and the structured synthesis of qualitative material into design artefacts such as personas, scenarios, and candidate system functionalities. Said design artefacts were then subsequently deployed in hands-on Co-design sessions with end-users, enabling collaborative exploration of system design possibilities.

LLM-supported analysis was conducted using TALLMesh, a bespoke qualitative analysis application developed to support human-controlled thematic analysis (De Paoli & Fawzi, 2025). TALLMesh provides a graphical user interface for interacting with LLMs during specific phases of thematic analysis, enabling researchers to execute structured prompts, control model parameters, and inspect and revise generated outputs. The tool is explicitly designed to support thematic analysis without automating interpretation or decision-making. It thus provides advanced LLM-based support whilst still allowing qualitative researchers to maintain key philosophical and epistemological assumptions during user-centred design.

Within this workflow, the LLM does not function as an analyst or collaborator. All substantive analytic work, including framing the analysis, selecting and refining themes, correcting inaccuracies, and validating outputs against raw data, is carried out by researchers with expertise in qualitative methods and close familiarity with the interview material. This position aligns with the qualitative nature of Co-design (De Paoli, 2025).

## WHAT IS IT GOOD FOR?

The workflow is particularly suited to design-oriented qualitative research, and analysis of Co-creation qualitative data, where thematic analysis is used to translate empirical insights into artefacts that support collaborative design and innovation processes.

In practice, qualitative analysis in design contexts is often conducted under significant time constraints, as design processes move rapidly from research to synthesis and decision-making. This pace makes the substantial volume of repetitive analytic work, such as summarising interview data, organising patterns across participants, and producing consistent representations of user needs, particularly challenging to manage. Maintaining progress within project timelines and sustaining analytic momentum is a central concern, and LLM-supported processing can reduce the time required for these tasks. Used in this way, LLMs allow researchers to focus more fully on interpretation, sense-checking, and discussion. Prior work has shown that LLMs can productively assist with aspects of qualitative synthesis when outputs are treated as provisional and subject to expert review rather than accepted as analytic conclusions (De Paoli, 2023; Barany et al., 2024; Yan et al., 2024).

Co-creation in this workflow, therefore, occurs among human participants, supported by computational tools. Intermediate artefacts such as codes, themes, personas, and scenarios act as shared reference points that enable collaborative sense-making across researchers, designers, developers, and project partners. This arrangement is consistent with accounts of joint action that emphasise coordinated human roles and shared responsibility without assuming equal or symmetrical contributions (Pacherie, 2012).

The workflow is most useful in settings where:

- Co-design teams are working with strict time constraints, but do not want to abandon a rigorous qualitative approach,
- researchers have strong familiarity with qualitative data and thematic analysis,
- thematic analysis is used to inform design or innovation decisions rather than to interpret deeply personal or therapeutic experiences, and
- analytic outputs must support collaborative sense-making across professional roles.

It is not intended as a general solution for all forms of qualitative analysis, nor as a substitute for qualitative expertise.

## HOW TO USE IT

The workflow follows the broad logic of reflexive thematic analysis (Braun & Clarke, 2006), with LLM processing introduced at clearly defined and controlled points when analysing user interview data. Figure 1 summarises the workflow, highlighting the key stages where LLM processing, researcher expertise and judgement, and wider stakeholder engagement are required.

First, researchers engage in familiarisation with user interview data and define the analytic focus in relation to project goals. This stage establishes the interpretive frame for the analysis and clarifies how insights will be translated into design artefacts.

Second, researchers prepare the dataset and set up a TALLMesh project. TALLMesh provides a visual interface for uploading and organising interview transcripts, managing project folders, and tracking intermediate analytic outputs. The application includes guidance pages that outline key phases of thematic analysis and clarify where LLM support is intended to assist rather than replace human judgement (De Paoli & Fawzi, 2025).

Third, within TALLMesh, researchers configure parameters for LLM processing. This includes selecting the LLM model, choosing from preset prompts aligned with thematic analysis tasks, or defining custom prompts tailored to the analytic focus.

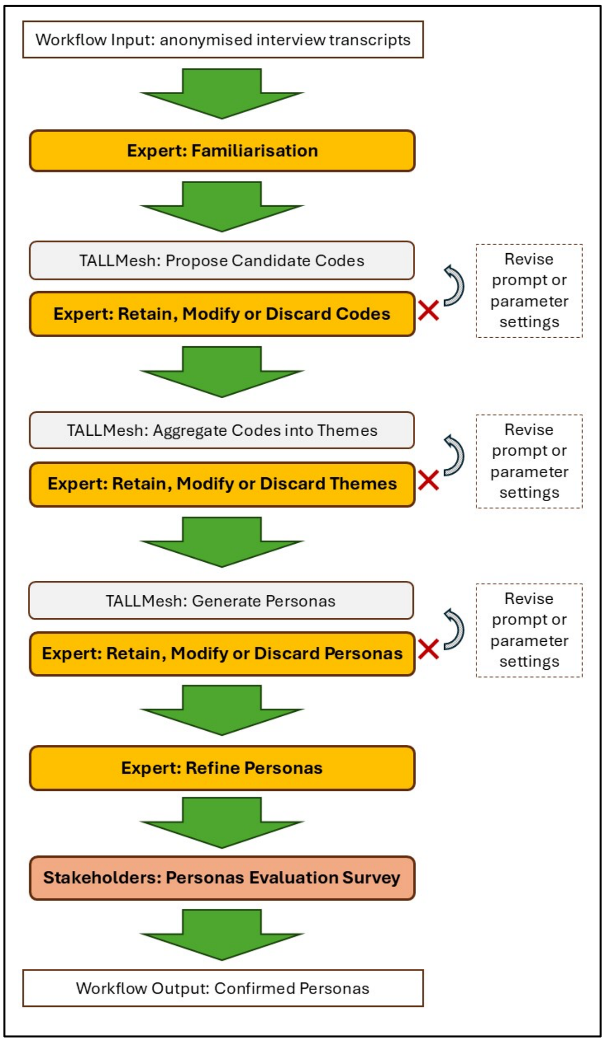


**Fig. 1**: AI-supported thematic analysis workflow highlighting key expert (yellow), TALLMesh (grey) and stakeholder (Orange) stages.

TALLMesh exposes controllable model parameters such as temperature (relating to output randomness, increasing or reducing predictability of the model) and top-p (governing probabilistic sampling), allowing researchers to prioritise reproducible, low-variance outputs. Setting these parameters conservatively supports replicability and limits creative divergence during pattern-processing tasks.

Fourth, TALLMesh supports interactive execution of key thematic analysis steps. Researchers can:

- generate initial codes that reflect design-relevant user insights (e.g., frustrations, goals), where the LLM proposes candidate codes linked to excerpts from interview transcripts;
- support theme generation, where sets of codes are aggregated into higher-level design-focused themes; and

- conduct code reduction and refinement, using dedicated interface views to remove duplicates, merge overlapping concepts, and adjust labels.

At each stage, outputs are displayed within the interface for inspection and editing. Researchers decide which codes or themes to retain, modify, or discard, and can iteratively rerun steps with revised prompts or parameter settings, remaining critical to every stage of the analysis. All outputs remain visible and traceable to source data.

Fifth, researchers iteratively compare LLM outputs with raw transcripts, using reflexive judgement to assess accuracy, relevance, and loss of nuance. Analytic decisions are discussed within the research team, reinforcing shared responsibility and collective sense-making (Braun & Clarke, 2021).

Finally, analytic artefacts are validated through collaborative review and wider stakeholder engagement. In the originating study, this included structured evaluation of personas by project partners and collective agreement on which artefacts would be taken forward. This step extends Co-creation beyond the core research team while maintaining clear boundaries around analytic responsibility. Stakeholders contribute to evaluation and selection, but interpretive accountability remains with the researchers.

### Scope and Limitations

Effective use of this workflow depends on researcher expertise and close engagement with the data. It is not suitable for use by analysts without training in qualitative methods, nor for contexts where interpretation of deeply personal, emotive, or therapeutic material is central.

While the workflow supports Co-creation through shared sense-making and collaborative evaluation, it does not distribute interpretive responsibility evenly across all participants or extend agency to LLMs. Interpretive accountability remains with researchers, reflecting ethical and methodological commitments in qualitative research.

In sum, responsible use of the workflow depends on explicit governance of analytic decisions and continued development of qualitative competence, such that LLM support remains bounded and accountable.

## ACKOWLEDGEMENTS

This work was supported by the European Union under Grant Agreement No. 101187940 (project LUMEN) and Grant Agreement No. 101188018 (project GRAPHIA). The views expressed are solely those of the authors and do not necessarily reflect those of the European Union.

## CONFLICTS OF INTEREST

None to declare.

## REFERENCES

Barany, A. et al. (2024). ChatGPT for Education Research: Exploring the Potential of Large Language Models for Qualitative Codebook Development. In: Olney, A.M., Chounta, IA., Liu, Z., Santos, O.C., Bittencourt, I.I. (eds) *Artificial Intelligence in Education. AIED 2024*. Lecture Notes in Computer Science, vol 14830. Springer, Cham. https://doi.org/10.1007/978-3-031-64299-9_10

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2): 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2021). Can I use TA? Should I use TA? Comparing reflexive thematic analysis and other pattern-based analytic approaches. *Counselling and Psychotherapy Research,* 21(1): 37–47. https://doi.org/10.1002/capr.12360

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(4): 997-1019. https://doi.org/10.1177/08944393231220483

De Paoli, S. (2025). Why I Reject the Use of Generative Artificial Intelligence in Qualitative Analysis. Available at SSRN: https://ssrn.com/abstract=5722002 or http://dx.doi.org/10.2139/ssrn.5722002

De Paoli, S., & Fawzi, A. (2025). TALLMesh: a simple application for performing thematic analysis with large language models. arXiv:2504.13892. https://doi.org/10.48550/arXiv.2504.13892

Morrow, D., Pehar, F., Silver, C. A., Fletcher, K. A., & De Paoli, S. (2025). Towards Human-AI collaborative methods in UX: a qualitative case study for discovery research platforms*. In 38th International BCS Human-Computer Interaction Conference* (pp. 519-529). BCS Learning & Development. https://doi.org/10.14236/ewic/BCSHCI2025.58

Pacherie, E. (2012). *The Phenomenology of Joint Action: Self-Agency versus Joint Agency*. In: Joint Attention, 343–390. The MIT Press. https://doi.org/10.7551/mitpress/8841.003.0017

Pearce, T., Maple, M., Shakeshaft, A., Wayland, S., & McKay, K. (2020). What is the co-creation of new knowledge? A content analysis and proposed definition for health interventions. *International Journal of Environmental Research and Public Health, 17*(7), 2229. https://doi.org/10.3390/ijerph17072229

Yan, L., Echeverria, V., Fernandez-Nieto, G. M., Jin, Y., Swiecki, Z., Zhao, L., Gašević, D., & Martinez-Maldonado, R. (2024). Human-AI Collaboration in Thematic Analysis using ChatGPT: A User Study and Design Recommendations. *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 1–7. https://doi.org/10.1145/3613905.3650732