DOI: https://doi.org/10.56198/5M1RHFUMB

Tutorial–Authoring a Personal GPT for Your Research and Practice: How We Created the QUAL-E Immersive Learning Thematic Analysis Helper

Leonel Morgado[1] and Dennis Beck[2]

[1] LE@D, CIAC, CEG, Universidade Aberta, Coimbra, Portugal & INESC TEC, Porto, Portugal
[2] University of Arkansas, Fayetteville, AR, U.S.A.
leonel.morgado@uab.pt , debeck@uark.edu

Abstract. Thematic analysis in qualitative research is a time-consuming and systematic task, typically done using teams. Team members must ground their activities on common understandings of the major concepts underlying the thematic analysis, and define criteria for its development. However, conceptual misunderstandings, equivocations, and lack of adherence to criteria are challenges to the quality and speed of this process. Given the distributed and uncertain nature of this process, we wondered if the tasks in thematic analysis could be supported by readily available artificial intelligence chatbots. Our early efforts point to potential benefits: not just saving time in the coding process but better adherence to criteria and grounding, by increasing triangulation between humans and artificial intelligence. This tutorial will provide a description and demonstration of the process we followed, as two academic researchers, to develop a custom ChatGPT to assist with qualitative coding in the thematic data analysis process of immersive learning accounts in a survey of the academic literature: *QUAL-E Immersive Learning Thematic Analysis Helper*. In the hands-on time, participants will try out QUAL-E and develop their ideas for their own qualitative coding ChatGPT. Participants that have the paid ChatGPT Plus subscription can create a draft of their assistants. The organizers will provide course materials and slide deck that participants will be able to utilize to continue development of their custom GPT. The paid subscription to ChatGPT Plus is not required to participate in this workshop, just for trying out personal GPTs during it.

Keywords: ChatGPT, GPT, Qualitative Analysis, Thematic Analysis.

1 Introduction and Background

A critical aspect of qualitative research is conducting thematic analysis. This involves a time-consuming process involving selection, extraction and assigning of meaning to relevant text excerpts. A crucial step of this process is the coding of the materials under analysis. This is the process where researchers (e.g. coders) apply their grounding and criteria to the materials, either as a whole or segmented, to generate codes, which basically are descriptive labels that surmise the main piece of information in the materials [1]. Given that people will experience subtle differences or equivocal interpretations in their grounding, and varying levels of adherence to coding criteria, the quality of the process is improved through extended involvement of multiple coders: triangulation. Triangulation is the use of multiple methods or data sources in qualitative research to develop a comprehensive understanding of phenomena [2]. It also has been viewed as a qualitative research strategy to test validity through the convergence of information from different sources. Thus, coding can be a weak point in the thematic analysis process because of insufficient number of coders, insufficiently trained coders, or simply due to differences in understanding and application of grounding concepts and coding criteria.

We sought to experiment with using an artificial intelligence variety, called Large Language Models (LLMs), to address these shortcomings through a publicly-available chatbot interface, ChatGPT. From its inception, artificial intelligence has made significant breakthroughs beyond answering standard search engine queries toward

analysis and evaluation of data [3]. In 2022, The Midjourney and ChatGPT chatbot-based interfaces enabled the public to access modern artificial intelligence models, for image generation and text generation, respectively, launching the current wave of enthusiasm with these technologies and raising awareness of their potential [4, 5].

The underlying language models vary immensely in their level of quality and abilities, with – at the time of writing – analyses pointing out towards OpenAI's GPT-4 and Google's Gemini Ultra being the most capable [6]. Since at the time of writing only the former is available to the public and enables personal customization and personal data content, we have employed GPT-4 for our work and it will be used for this workshop: we called our GPT "QUAL-E Immersive Learning Thematic Analysis Helper", and it is publicly available (link in the references) [7]. OpenAI's ChatGPT Plus service allows subscribing users a simple interface to develop their personal AI assistants based on GPT-4. The tool also supports improved function-calling for 3rd party tools.

The rationale for this use in support of qualitative thematic analysis coding is that LLMs have shown remarkable good quality in a variety of tasks involving interpretation of textual content. A recent analysis revealed that several LLMs "not only achieve high accuracy overall, but also capture fine-grained variation in human linguistic judgments", and that "GPT-3.5 Turbo and GPT-4 outperform humans according to DGL's normative grammaticality coding", concluding that "LLMs show strong and human-like grammatical generalization capabilities" [8]. Qualitative, thematic analysis requires team members to reach a common understanding of what the material under analysis means, and how it can be reflected in a code – and also what that means and how it relates to other codes. Thus, linguistic judgments, a task that LLMs are good at, are at the heart of the coding process. Thus, there was some expectation that LLMs might help lessen misunderstandings, equivocations, and lack of adherence to criteria, all challenges to the quality and speed of this process. Given the distributed and uncertain nature of this process, we further wondered if the coding tasks in thematic analysis could be distributed by running multiple instances of artificial intelligence chatbots, thus improving not only linguistic judgments, but also triangulation.

2 Workshop Description

This two-hour workshop will cover five main priorities:

1. An overview of our motivation and purpose in creating QUAL-E Immersive Learning Thematic Analysis Helper.
2. A description of the step-by-step iterative design process we utilized in creating QUAL-E.
3. A demonstration of QUAL-E operation and abilities.
4. Hands-on time where the researchers leading the workshop will help individual participants from the ideation phase, assembly of materials, and crafting of instructions.
5. Participants with a paid ChatGPT Plus account will have the opportunity to implement their personal GPT thematic coding helper and try it.

An overview of our motivation and purpose in creating QUAL-E Immersive Learning Thematic Analysis Helper can be found above. We followed a step-by-step iterative process to create the QUAL-E Immersive Learning Thematic Analysis Helper. We began by making a list of all of the tasks we wanted QUAL-E to be able to accomplish. We wanted QUAL-E to identify the kind of educational use discussed in each text excerpt. Thus, the first task was to train our AI assistant to identify a discrete concept within a specific text extraction (usually one to two sentences), and then to label that concept with a single word or short phrase that summarized the meaning of the educational use.

We accomplished these tasks by uploading several, peer reviewed articles and books that defined important vocabulary terms (e.g. immersion, "educational use", etc.). We also uploaded articles that we utilized in training ourselves in the qualitative coding process, directing ChatGPT to specific sections in those articles to reference how to perform thematic coding. Finally, ChatGPT 4.0 Plus allows the user to write instructions up to a limit of 8000 characters. Thus, we authored instructions which included 1) specific definitions to use and from which document in the knowledge base; 2) how to handle specific situations in the coding process and which sequence to follow in the coding process; 3) Any exceptions to that coding process; and 4) directions on the specific output structure required for the coding process.

Our early efforts point to potential benefits: not just saving time in the coding process but rather better adherence to criteria and grounding, by increasing triangulation between humans and artificial intelligence. (Triangulation is a qualitative research strategy to test validity through the convergence of data analysis from different sources). This workshop will provide a description and demonstration of the process we followed to develop a custom ChatGPT to assist with qualitative coding in the thematic data analysis process of immersive learning

accounts in a survey of the academic literature: QUAL-E Immersive Learning Thematic Analysis Helper. In the hands-on time, participants will try out QUAL-E and develop their ideas for their own qualitative coding ChatGPT. Participants that have the paid ChatGPT Plus subscription can create a draft of their assistants. The organizers will provide course materials and slide deck that participants will be able to utilize to continue development of their custom GPT. The paid subscription to ChatGPT Plus is not required to participate in this workshop, just for trying out personal GPTs during it.

The QUAL-E Immersive Learning Thematic Analysis Helper Workshop will help to train immersive learning researchers and practitioners in how to create their own custom GPT assistants. Equipped with this knowledge, the participants will be able to produce higher quality work and improve their own academic and professional contexts.

References

1. Green, J., Willis, K., Hughes, E., Small, R., Welch, N., Gibbs, L., Daly, J.: Generating best evidence from qualitative research: the role of data analysis. Australian and New Zealand Journal of Public Health. 31, 545–550 (2007). https://doi.org/10.1111/j.1753-6405.2007.00141.x
2. Patton, M.Q.: Enhancing the quality and credibility of qualitative analysis. Health Serv Res. 34, 1189–1208 (1999)
3. Wang, H., Fu, T., Du, Y., Gao, W., Huang, K., Liu, Z., Chandak, P., Liu, S., Van Katwyk, P., Deac, A., Anandkumar, A., Bergen, K., Gomes, C.P., Ho, S., Kohli, P., Lasenby, J., Leskovec, J., Liu, T.-Y., Manrai, A., Marks, D., Ramsundar, B., Song, L., Sun, J., Tang, J., Veličković, P., Welling, M., Zhang, L., Coley, C.W., Bengio, Y., Zitnik, M.: Scientific discovery in the age of artificial intelligence. Nature. 620, 47–60 (2023). https://doi.org/10.1038/s41586-023-06221-2
4. Severinson, J.: How far Midjourney has come in just a year, https://www.linkedin.com/pulse/how-far-midjourney-has-come-just-year-john-severinson/
5. Mesko, B.: The ChatGPT (Generative Artificial Intelligence) Revolution Has Made Artificial Intelligence Approachable for Medical Professionals. J Med Internet Res. 25, e48392 (2023). https://doi.org/10.2196/48392
6. Gemini Team, A Family of Highly Capable Multimodal Models, http://arxiv.org/abs/2312.11805, (2023)
7. Leonel Morgado, Dennis Beck: QUAL-E Immersive Learning Thematic Analysis Helper, https://chat.openai.com/g/g-LcesDN88c-qual-e-immersive-learning-thematic-analysis-helper
8. Hu, J., Mahowald, K., Lupyan, G., Ivanova, A., Levy, R.: Language models align with human judgments on key grammatical constructions, http://arxiv.org/abs/2402.01676, (2024)