CAQDAS, and Artificial Intelligence: New Challenges and Possibilities for Teaching Qualitative Data Coding

Nelson Paulus S[1] [0000-0002-8524-5167]
Byron Duhalde V[2] [0000-0003-0235-8623]
[1,2]Universidad de Santiago de Chile (USACH)
nelson.paulus@usach.cl

Abstract

The research evaluates the value and power of an AI tool such as Chat GPT as CAQDAS (Computer-Aided Qualitative Data Analysis Software) An experiment carried out using version 4.0 of the application is presented, using it as a standalone tool, even though traditional CAQDAS programs, such as ATLAS ti already announce a collaboration interface with Chat GPT. The focus is on the function of Open Coding, as one of the initial stages of the processing of qualitative material (and one of the most tiring) according to the original proposal of the Grounded Theory (GT) and which was adopted, due to their usefulness, for various other less canonical forms of qualitative analysis. An experiment is conducted with a sample of qualitative material associated with news related to migration issues. From this, through an inductive exercise facilitated by AI, a set of codes is obtained to give way to more complex instances of the analytical process (Axial Coding, Theoretical Elaboration, etc.). The existing potential in the use of the computational tool is revealed, in terms of power for a first (proposal) of Open Coding with optimization of human resources and time.

KEYWORDS: CAQDAS – ARTIFICIAL INTELLIGENCE – CODING – GROUNDED THEORY

Introduction

CAQDAS (Computer-Aided Qualitative Data Analysis Software) is used to designate a set of programs whose function is to assist in the analysis of non-numerable data in the social sciences. Among this type of program, one of the most popular is ATLAS-ti , which since 2023 indicates that it complements its performance with Chat-GPT's Open AI. With this interface, the software developers point out that substantial amounts of time can be saved in the tedious task of coding through coding with Artificial Intelligence (AI Coding).

However, the most striking thing about the latest advances in AI and CAQDAS is that on its own (and independently of its association with ATLAS-ti or other programs) Chat GPT 4.0 can support coding, suggesting new possibilities for the social sciences, suggesting strategies and systems unthinkable for the exploitation of qualitative material. The above applies, for example, to that stage of the first reduction of qualitative material, according to the canon of the Grounded Theory, referred to as Open Coding. The present work aims to estimate the scope of the challenges thus posed through an experiment whose horizon is to evaluate the potential of a tool such as Chat GPT 4.0 to support

the qualitative analysis process, even autonomously to the interface it may have with traditional qualitative analysis programs such as ATLAS-ti.

Theoretical Framework

Regarding the CAQDAS and for teaching purposes, we understand qualitative analysis as the set of "manipulations, transformations, operations, reflections and verifications that are carried out on the data in order to extract a relevant meaning in relation to the research problem." (Rodríguez, Gil and García, 1999; p 200)

Given the wide availability of secondary material to be exploited that we currently have on the Internet, after the indispensable prior instance of a first review, selection and cleaning that can be made of the data collected, what usually proceeds in the analysis is the reduction of complexity and the elaboration of order and meaning of the material via coding (Miles, Huberman, & Saldaña, 2014). For this work, we have focused exclusively on the initial stage of Open Coding which, with an inductive coding dynamic very typical of Grounded Theory (Requena, Planes & Miras, 2014; Andreu, García, & Pérez, 2007), and widely used by analytical approaches to qualitative material that are less canonical than the traditional GT, such as, among others, the interpretive and constructivist versions of the GT (see Sebastian, 2019) Initiatives that are very much in tune with what is presented here have been developed in recent years. In this regard, see Lopezosa, et al (2023); Lopezosa & Codina (2023); Lopezosa, Codina & Boté-Vericad (2023); Nikolopoulou, (2024); NguyenTrung, K. (2024) Goyanes, Lopezosa, & Jordá (2024). Safaei, & Longo, (2024).

Method

We carried out an experiment trying to reproduce a strictly inductive dynamic. To this end, a 300-page Word document was processed that synthesized news on migration from Chilean national journalistic media between June 1, 2023 and July 31, 2023. The exercise was carried out with Chat GPT 4.0 (paid version).

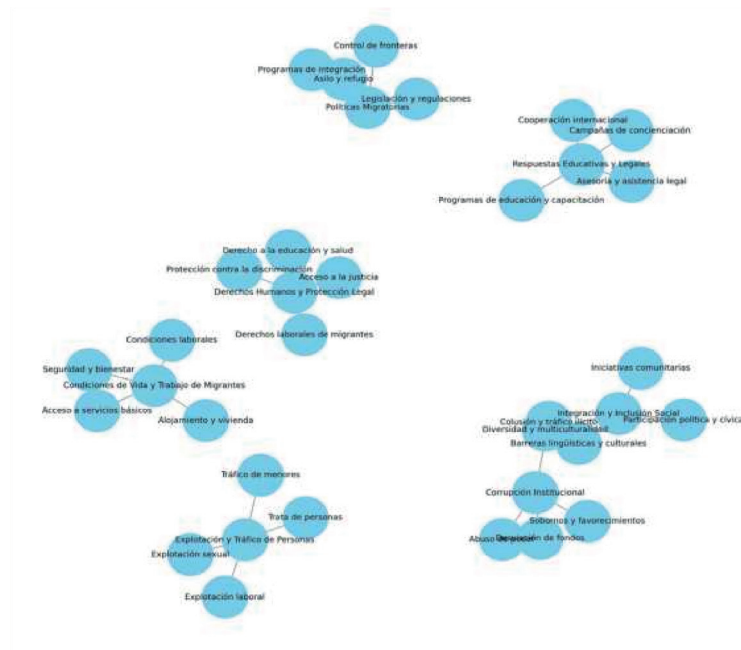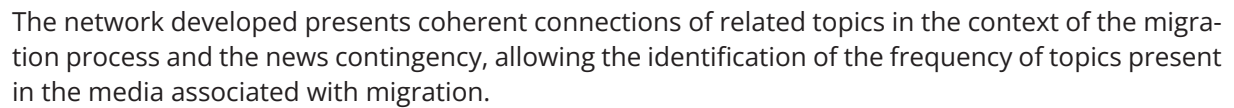**Figure 1: Result of the First Network.**

Figure nº2. Frequently Topic Network Output

The network developed presents coherent connections of related topics in the context of the migration process and the news contingency, allowing the identification of the frequency of topics present in the media associated with migration.

Discussion and Conclusions

When teaching CAQDAS, it has been usual for teachers to indicate that it is not the software that performs the analysis, but the analyst with the help of the respective software, since it assumes that there is an analytical strategy that does not go through the program but through the researcher who uses it. We believe that with AI tools such as Chat GPT 4.0, the idea that "it is not the program that performs the analysis" deserves to be reviewed in its most precise scope.

Limitations and Projections

For the moment, what has been experienced refers to the basic strategy of Open Coding, which is the first step in reducing material. Of course, it also advances, at least in terms of graphic representation in what could be Axial Coding, a subsequent procedure in the context of the GT canon. (Requena, Planes y Miras, 2014; Andreu, García, & Pérez, 2007). However, prudence indicates that we must properly weigh this first step associated with Open Coding with AI (Bryant, 2023), before automatically proceeding with thoughtless automation of GT procedures.

References

Andreu, J., García, A., & Pérez, A. (2007). Evolución de la teoría fundamentada como técnica de análisis cualitativo. Madrid. Centro de investigaciones sociológicas. Cuadernos Metodológicos, 40.

Bryant, A. (2023, June). AI Chatbots: Threat or Opportunity?. In Informatics (Vol. 10, No. 2, p. 49). MDPI.

Carrero, V., Soriano, R. M., & Trinidad, A. (2012). Teoría fundamentada Grounded Theory (Vol. 37). CIS.

Goyanes, M., Lopezosa, C., & Jordá, B. (2024). Thematic Analysis of Interview Data with ChatGPT: Designing and Testing a Reliable Research Protocol for Qualitative Research.

Lopezosa, C., & Codina, L. (2023). ChatGPT y software CAQDAS para el análisis cualitativo de entrevistas: pasos para combinar la inteligencia artificial de OpenAI con ATLAS. ti, Nvivo y MAXQDA.

Lopezosa, C., Codina, L., & Boté-Vericad, J. J. (2023). Testeando ATLAS. ti con OpenAI: hacia un nuevo paradigma para el análisis cualitativo de entrevistas con inteligencia artificial.

Lopezosa, C., Codina, L., Pont-Sorribes, C., & Vállez, M. (2023). Use of generative artificial intelligence in the training of journalists: challenges, uses and training proposal. Profesional de la información/Information Professional, 32(4).

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). Qualitative data analysis: A methods sourcebook.3rd.

Nguyen-Trung, K. (2024). ChatGPT in Thematic Analysis: Can AI become a research assistant in qualitative research? OSF Preprint.

Nikolopoulou, K. (2024). Generative Artificial Intelligence in Higher Education: Exploring Ways of Harnessing Pedagogical Practices with the Assistance of ChatGPT. International Journal of Changes in Education, 1(2), 103-111.

Rodríguez Gómez, G., Gil Flores, J., & García Jiménez, E. (1999). Metodología de la investigación cualitativa. Segunda edición. Aljibe

Safaei, M., & Longo, J. (2024). The End of the Policy Analyst? Testing the Capability of Artificial Intelligence to Generate Plausible, Persuasive, and Useful Policy Analysis. Digital Government: Research and Practice, 5(1), 1-35.

Sebastian, K. (2019). Distinguishing between the strains grounded theory: Classical, interpretive and constructivist. Journal for Social Thought, 3(1).