

CAQDAS, e Inteligencia Artificial: Nuevos desafíos y posibilidades para la enseñanza de la codificación de datos cualitativos

Nelson Paulus S.¹ [0000-0002-8524-5167]

Byron Duhalde V.² [0000-0003-0235-8623]

^{1,2}Departamento de Estudios Políticos

nelson.paulus@usach.cl

Resumen

La investigación evalúa el valor y la potencia de una herramienta de IA como Chat GPT como CAQDAS (Computer-Aided Qualitative Data Analysis Software). Se presenta un experimento realizado utilizando la versión 4.0 de la aplicación, utilizándola como herramienta autónoma, a pesar de que programas tradicionales de CAQDAS, tales como ATLAS ti ya anuncian una interfaz de colaboración con Chat GPT. Se pone el foco en la función de Codificación Abierta, como una de las etapas iniciales del procesamiento de material cualitativo (y una de las más fatigosas) conforme a la propuesta original de la Grounded Theory (GT) y que adoptan, por su utilidad, diversas otras formas menos canónicas de análisis cualitativo. Se realiza un experimento con una muestra de material cualitativo asociado a noticias referidas a temas migratorios. A partir de esto se obtiene, a través de un ejercicio inductivo facilitado por la IA, un conjunto de códigos dispuestos para dar paso a instancias más complejas del proceso analítico (Codificación Axial, Elaboración Teó-rica, etc.). Se releva el potencial existente en la utilización de la herramienta computacional, en términos de potencia para una primera (propuesta) de Codificación Abierta con optimización de recurso humano y de tiempo.

PALABRAS CLAVE: CAQDAS – INTELIGENCIA ARTIFICIAL – CODIFICACIÓN – GROUNDED THEORY

Introducción

Con CAQDAS (Computer-Aided Qualitative Data Analysis Software) se designa el conjunto de programas cuya función es ayudar el análisis de datos no numerizables en las ciencias sociales. Entre este tipo de programas, uno de los más populares es ATLAS-ti, el que desde 2023 indica que complementa su desempeño con Open AI de Chat-GPT. Con esta interfaz los desarrolladores del software señalan que se podrán ahorrar sustantivas cantidades de tiempo en la tediosa labor de codificación a través de la codificación con Inteligencia Artificial (AI Coding).

Sin embargo, lo más llamativo en los últimos avances en materia de IA y CAQDAS es que por sí solo (y de manera autónoma a su asociación a ATLAS-ti u otros programas) Chat GPT 4.0 puede apoyar en la codificación, sugiriendo nuevas posibilidades para las ciencias sociales, sugiriendo estrategias y sistemas impensados para la explotación del material cualitativo. Valga lo anterior, por ejemplo, para aquella etapa de primera reducción de material cualitativo, según el canon de la Grounded Theory, referida como Codificación Abierta.

El presente trabajo pretende estimar en sus alcances los desafíos así planteados a través de un experimento cuyo horizonte es evaluar el potencial de una herramienta como Chat GPT 4.0 para apoyar el proceso de análisis cualitativo, incluso de manera autónoma a la interfaz que pueda tener con programas tradicionales de análisis cualitativo como ATLAS-ti.

Marco Teórico

En cuanto al CAQDAS y para efectos de la docencia, entendemos por análisis cualitativo el conjunto de “manipulaciones, transformaciones, operaciones, reflexiones y comprobaciones que se realizan sobre los datos con el objeto de extraer un significado relevante en relación a la problemática de investigación” (Rodríguez, Gil y García, 1999; p 200).

Dada la amplia disponibilidad de material secundario a ser explotado del que actualmente disponemos en virtud de Internet, tras la indispensable instancia previa de una primera revisión, selección y limpieza que se pueda hacer de los datos recolectados, lo que procede usualmente en el análisis es la reducción de complejidad y confección de orden y de sentido del material vía codificación (Miles, Huberman, y Saldaña, 2014).

Para efectos del presente trabajo nos hemos centrado de manera exclusiva en la etapa inicial de la Codificación Abierta en la cual, con una dinámica de codificación inductiva muy propia de la Grounded Theory (Requena, Planes y Miras, 2014; Andreu, García y Pérez, 2007), y muy utilizada por aproximaciones analíticas de material cualitativo menos canónicas que la GT tradicional, como son, entre otras, las versiones interpretativa y la constructivista de la GT (ver Sebastian, 2019) Iniciativas muy en sintonía con lo aquí presentado han sido desarrolladas en los últimos años. Al respecto, ver Lopezosa, et al (2023); Lopezosa & Codina (2023); Lopezosa, Codina & Boté-Vericad (2023); Nikolopoulou (2024); Nguyen-Trung, K. (2024) Goyanes, Lopezosa, & Jordá (2024). Safaei, & Longo, (2024).

Método

Realizamos un experimento intentando reproducir una dinámica estrictamente inductiva. Para ello se procesó un documento de Word de 300 páginas que sintetizaba noticias sobre migración provenientes de medios periodísticos nacionales chilenos entre el 1 de junio de 2023 y 31 de julio de 2023. El ejercicio se realizó con Chat GPT 4.0 (versión de pago)

Resultados

Primero se solicitó a la IA la identificación de códigos y subcódigos basados en los temas recurrentes en las noticias, obteniéndose por parte de la aplicación siete códigos principales:

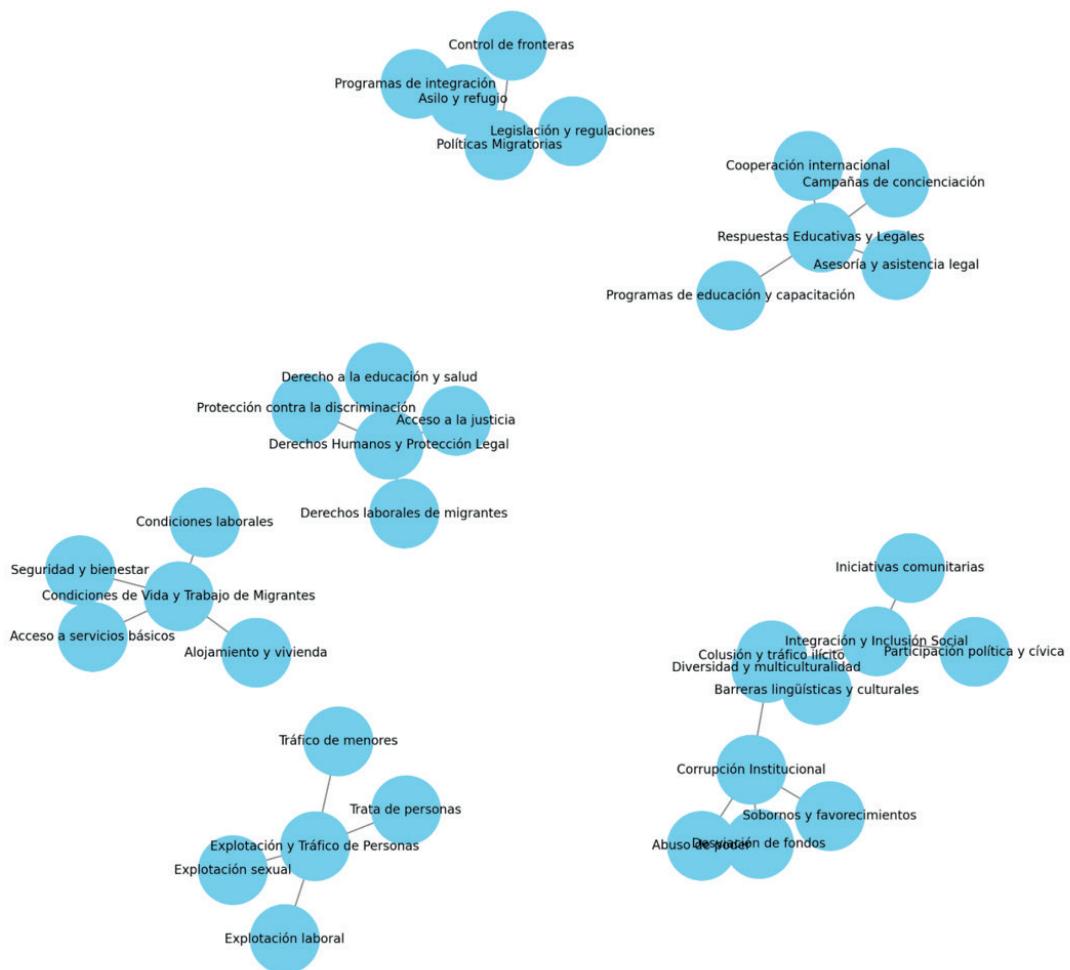
- Corrupción Institucional,
- Políticas Migratorias,
- Derechos Humanos y Protección Legal,
- Condiciones de Vida y Trabajo de Migrantes,
- Explotación y Tráfico de Personas,
- Respuestas Educativas y Legales,
- Integración y Inclusión Social.

* Para cada código, además, se identificaron subcódigos específicos que detallan aspectos particulares del tema principal.

Posteriormente a ello, se solicitó la creación de una red que vinculara los códigos con los subcódigos. La IA utilizó la biblioteca NetworkX⁶ para construir una red que representara la relación entre códigos y subcódigos.

Los códigos principales se añadieron como nodos, y se conectaron a sus respectivos subcódigos mediante aristas. Además, se añadieron conexiones entre subcódigos de diferentes códigos principales para reflejar relaciones temáticas adicionales identificadas durante el análisis.

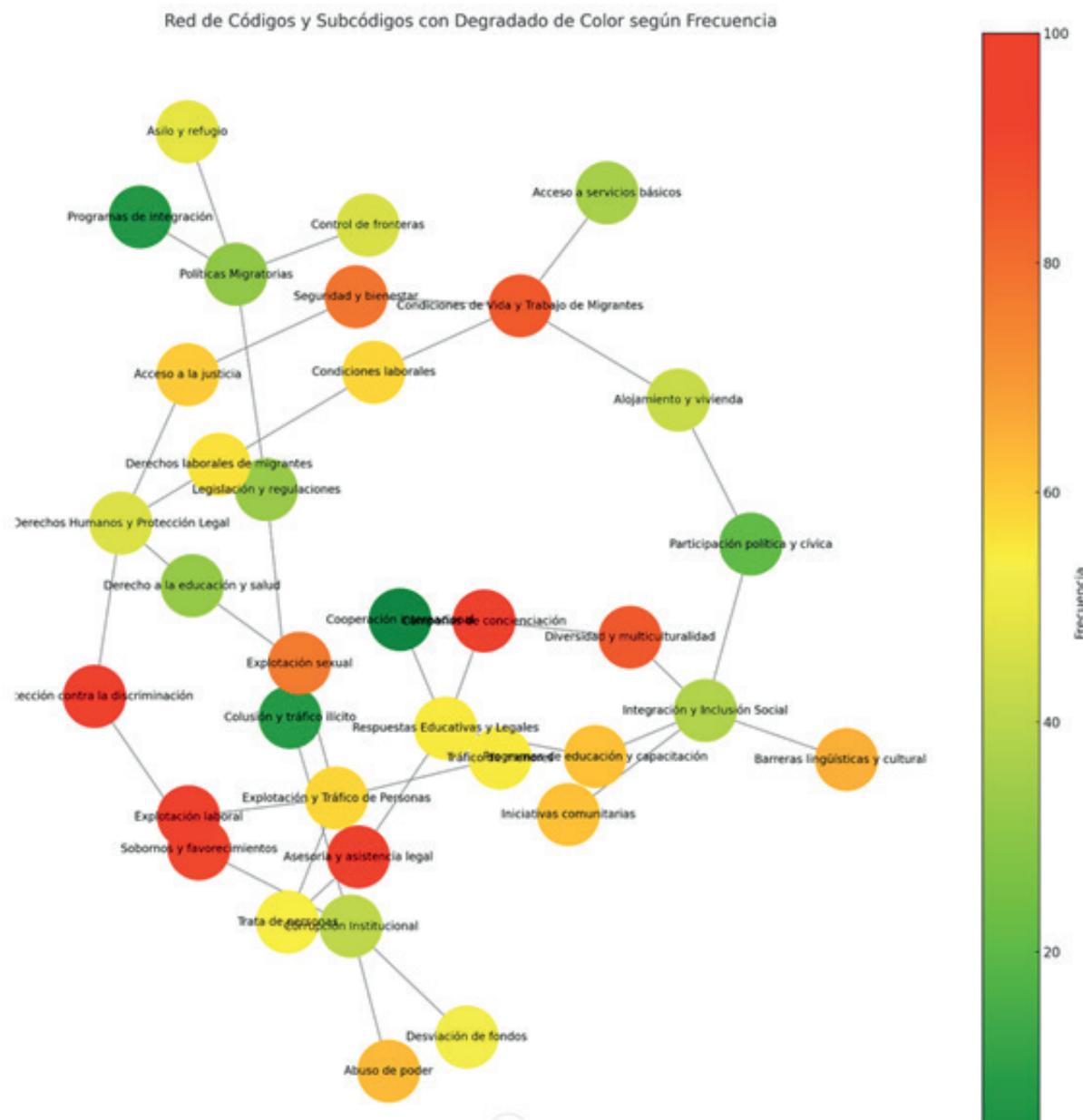
Figura n.º 1. Resultado de la Primera Red



Para precisar la interpretación de los temas identificados, se solicitó a la IA en función de la frecuencia de los temas asignar un degradado de colores desde verde (baja frecuencia) a rojo (alta frecuencia) para representar visualmente estas frecuencias. Las frecuencias se normalizaron a través de la biblioteca matplotlib.colors se utilizó para crear un mapa de colores lineal.

6 Network Analysis in Python (<https://networkx.org>) de acceso abierto.

Figura n.º 2. Resultado de red con frecuencia de temas



La red elaborada presenta conexiones coherentes de los temas relacionados en el contexto del proceso migratorio y la contingencia noticiosa, permitiendo identificar la frecuencia de temas presentes en los medios asociados a la migración.

Discusión y Conclusiones

Al enseñar CAQDAS, ha sido usual por parte de los docentes indicar que no es el software quien realiza el análisis, sino que el analista con ayuda del software respectivo, en tanto supone que hay una estrategia analítica que no pasa por el programa sino por el investigador que lo utiliza. Creemos que con herramientas de IA tales como Chat GPT 4.0, aquello de que “no es el programa el que realiza el análisis” merece ser revisado en su alcance más preciso.

Limitaciones y Proyecciones

De momento, lo experimentado se refiere a la estrategia básica de Codificación Abierta, lo cual es un primer paso de reducción de material. Por supuesto también avanza, al menos en términos de representación gráfica en aquello que podría ser la Codificación Axial, procedimiento subsecuente en el contexto de los señala el canon de la GT. (Requena, Planes y Miras, 2014; Andreu, García y Pérez, 2007). No obstante la prudencia indica que debemos ponderar adecuadamente este primer paso asociado a la Codificación Abierta con IA (Bryant, 2023), antes de proseguir de manera automática con un irreflexivo automatización de los procedimientos de la GT.

Referencias

- 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).