

Exploración metodológica del análisis temático de datos cualitativo con herramientas de IA generativa

Mercedes Reguant-Álvarez ^a

Francesc Martínez-Olmo ^b

Universitat de Barcelona. España.

^a <https://orcid.org/0000-0002-0830-7854>

^b <https://orcid.org/0000-0002-9719-2808>

Resumen

Introducción. La Inteligencia Artificial Generativa (IAGen) introduce herramientas con potencial transformador en la investigación y ofrece nuevas posibilidades para la codificación, categorización y sistematización de información narrativa. El objetivo principal de este estudio es explorar metodológicamente la eficacia, la fiabilidad y las implicaciones éticas (particularmente la precisión en la extracción de citas y la minimización del sesgo) que se presentan al realizar un análisis cualitativo temático asistido por IAGen. **Método.** Se redactó un mismo *prompt* en tres herramientas de IAGen (*Copilot*, *DeepSeek* y *ChatGPT*) aplicado a un corpus textual correspondiente a la transcripción anonimizada de un grupo de discusión sobre la herramienta educativa *e-diario*. Se compararon los resultados obtenidos por cada herramienta. **Resultados.** Las tres herramientas aportaron toda la información solicitada en el *prompt* e identificaron entre 7 y 9 categorías temáticas, con una alta coincidencia entre ellas. **Conclusiones.** La IAGen se presenta como una herramienta útil para la fase exploratoria del análisis cualitativo, permitiendo una aproximación inicial automatizada a grandes volúmenes de datos. Sin embargo, su uso debe estar acompañado de supervisión humana. La IAGen puede complementar el análisis tradicional, siempre que se utilice de forma ética, crítica y responsable.

Palabras clave: inteligencia artificial, inteligencia artificial generativa, análisis temático, análisis cualitativo, ética de la tecnología, ética de la ciencia, método científico

Reguant-Álvarez, M., & Martínez-Olmo, F. (2025). Exploración metodológica del análisis temático de datos cualitativo con herramientas de IA generativa. *Zenodo*. <https://doi.org/10.5281/zenodo.17367698>

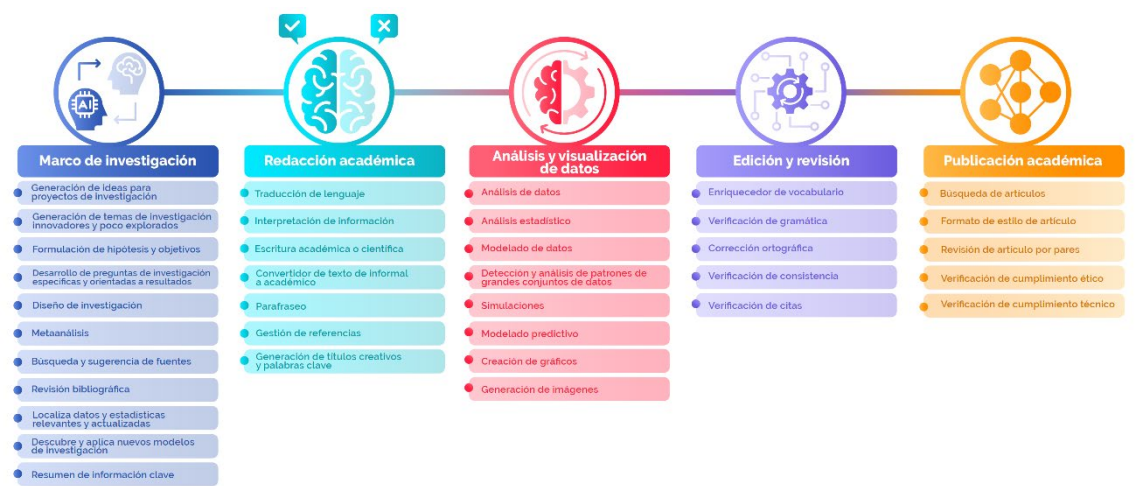
© 2025 Los autores. Este documento es de acceso abierto sujeto a la licencia Reconocimiento 4.0 Internacional de Creative Commons, la cual permite utilizar, distribuir y reproducir por cualquier medio sin restricciones siempre que se cite adecuadamente la obra original. Para ver una copia de esta licencia, visite <https://creativecommons.org/licenses/by/4.0/>



Introducción

En la actualidad, el uso de la Inteligencia Artificial Generativa (IAGen) —*Generative AI* (*GenAI*), en inglés—, irrumpe en todos los ámbitos y campos del saber; cada día se desarrollan nuevas aplicaciones y, por ello, es indispensable pensar en ella y sus posibles usos y potencialidades. El auge de la IAGen ha introducido herramientas con potencial transformador también en el campo de la investigación académica (ver Figura 1). La adopción de estas tecnologías es rápida, como lo demuestra la estimación de que el 1 % de los *abstracts* enviados al repositorio *arXiv* ya en 2023 contenían evidencia de su uso (Gray, 2024). Dentro del ciclo de investigación, las fases de procesamiento y análisis de datos se han beneficiado notablemente de la AIGen. Específicamente, estas herramientas se consideran valiosas para los procesos de codificación y categorización de datos, y poseen el potencial de apoyar significativamente el análisis cualitativo (*QDA*, *qualitative data analysis*) (Bjelobaba et al., 2025). Herramientas de Modelos de Lenguaje Grande (LLM), como ChatGPT, han emergido como instrumentos atractivos para el QDA, particularmente en las ciencias sociales, ya que ofrecen formas más rápidas de procesar datos y permiten ahorrar tiempo y recursos a los investigadores (Sandín Esteban & Sánchez Martí, 2024).

Figura 1. Usos generales de la IA en la investigación académica



Fuente: Zohery (2023), adaptado por Miguel Morales-Chan (@amoraleschan).

La IAGen es una rama de la IA que crea contenido nuevo, como texto, imágenes, música, audio y videos, a partir de datos existentes con los que se entrena. Se basa en modelos de aprendizaje profundo que aprenden patrones y relaciones en conjuntos de datos y luego los usan para generar nueva información (Cruz Argudo et al., 2024).

El análisis cualitativo se centra en la organización, interpretación y extracción de conclusiones significativas a partir de grandes volúmenes de datos no estructurados o narrativos. Hasta ahora, al enfrentar un volumen muy importante de datos cualitativos —ya sean grupos de discusión, preguntas abiertas de un cuestionario, noticias de prensa, intercambio en chats o cualquier otro tipo de relatos—, una forma de acercarse a la deconstrucción y reconstrucción de toda la información obtenida comporta un largo camino de lectura y relectura, segmentación, categorización y codificación.

Las capacidades de la IAGen resultan atractivas para este tipo de análisis, ya que pueden asistir en tareas como la generación de resúmenes automáticos, la identificación de temas, patrones y tendencias, la sugerencia de códigos iniciales, y la sistematización de la información en varios formatos. Sin embargo, la naturaleza probabilística de la IAGen, que predice el texto basándose en patrones estadísticos, implica que sus respuestas no conllevan una comprensión en el sentido humano, que involucra conciencia y razonamiento abstracto (Sandín Esteban & Sánchez Martí, 2024).

Con todo, la integración de IAGen en la investigación plantea importantes desafíos éticos que incluyen la falta de transparencia, el sesgo inherente a los datos de entrenamiento y la fabricación de contenido —fenómeno conocido como "alucinaciones" o información falsa presentada como factual— (Bjelobaba et al., 2025). En el contexto específico del análisis cualitativo, una preocupación ética importante es el riesgo de malinterpretación o fabricación de datos de investigación. Se ha documentado que las herramientas de IAGen pueden generar citas *ficticias* para ilustrar temas, incluso si estas citas coinciden con la temática identificada, lo que subraya la problemática de las *alucinaciones* y la falta de transparencia (Mahyoob

et al., 2023). De hecho, la dificultad para detectar contenido generado por IAGen es tan alta que incluso revisores expertos y plataformas antiplagio han fallado en su identificación (Odri & Yoonb, 2023). A pesar de que existen algunos estudios que han explorado el uso de IAGen para el análisis cualitativo (por ejemplo, Amani et al., 2025; Cook et al., 2025; Wheeler, 2025), la investigación sobre la aplicación de estas herramientas en el contexto más amplio de la investigación académica sigue siendo limitada. Además, la constante evolución de las plataformas de IAGen hace que una evaluación exhaustiva y continua, herramienta por herramienta, sea impracticable.

Dada la limitada investigación comparativa de modelos específicos de IAGen aplicados a tareas fundamentales de investigación, como el análisis cualitativo, y la necesidad de mitigar los riesgos de alucinación y sesgo, este estudio se enfoca en una evaluación práctica de tres modelos representativos de la IAGen utilizados en el análisis de un mismo corpus textual. Específicamente, se comparan *Copilot* (un modelo conocido por su utilidad en tareas de código y texto), *DeepSeek* (como ejemplo de una plataforma IAGEN emergente) y *ChatGPT* (un modelo dominante y de uso generalizado) (Van Noorden & Perkel, 2023). Al contrastar el desempeño de estas herramientas en la codificación y la identificación de temas, se busca aportar una perspectiva empírica sobre los patrones de riesgo y utilidad que surgen al integrar diversos modelos de IAGEN en las metodologías cualitativas.

El objetivo principal de este estudio es explorar metodológicamente la eficacia, la fiabilidad y las implicaciones éticas (particularmente la precisión en la extracción de citas y la minimización del sesgo) que se presentan al realizar un análisis cualitativo temático asistido por IAGen. A partir de esta evaluación, el estudio busca ofrecer orientaciones prácticas para el personal investigador que busca integrar estas herramientas de forma responsable, transparente y ética en sus procesos de análisis de datos cualitativos. Así, este documento no pretende exponer un estudio profundo sobre el origen, los sesgos, o el funcionamiento de las redes neuronales con las que se desarrolla la IAGEN, sino más bien una guía práctica a partir de una experiencia en su uso para el análisis de contenido temático sobre un volumen importante de narraciones.

Método

Para ciertos investigadores acostumbrados a trabajar con el análisis de textos, y más específicamente con el análisis temático, el contar con una visión general previa del relato puede resultar de mucha ayuda. Además, el hecho de que este conocimiento inicial del texto pueda ser totalmente inductivo y automatizado, hace del planteamiento de este artículo una herramienta eficaz.

En los siguientes párrafos, se explica la forma en que se procedió para obtener una primera aproximación a la transcripción de un grupo de discusión¹, utilizando las herramientas de IAGen.

El primer paso fue redactar el siguiente *prompt* y elegir tres modelos de IAGen utilizados con cierta frecuencia:

A continuación, voy a subir la transcripción de un grupo de discusión. El objetivo de este era valorar la herramienta e-diario, un diario virtual, compartido entre el alumnado de prácticas en la formación de maestros de primaria, de 3 universidades españolas. Me gustaría tener un análisis temático del grupo de discusión, en el que se aprecien las categorías, su frecuencia de aparición y algún ejemplo extraído del texto, en cada caso. También quiero un resumen del contenido del grupo de discusión. Y, un resumen breve, de no más de un párrafo del pensamiento de cada informante. Todo con un lenguaje académico.

Este mismo *prompt* se introdujo en *Copilot* —versión GPT-5 (Microsoft, 2025)—, *DeepSeek* —versión R1-0528 (DeepSeek, 2024)— y *ChatGPT* —versión GPT-5 (OpenAI, 2025)—, así como el texto del grupo de discusión previamente transcrito y anonimizado.

Una vez obtenidos los resultados de cada herramienta de IAGen, los autores compararon cada una de las respuestas aportadas, a fin de poder establecer los elementos clave del análisis: categorías, citas y resúmenes.

¹ La transcripción del grupo de discusión utilizado forma parte de la información recogida en el proyecto «E-Diari compartit: una experiència col·laborativa en xarxa en l'especialitat de mestres d'Educació Física de la Facultat de Formació del Professorat».

Resultados

El resultado fue interesante, ya que cada una de las herramientas ofreció informaciones similares, aunque con matices. Esto llevó a hacer, por parte de los autores, un análisis de las coincidencias y diferencias. A continuación, se muestran las categorías creadas por cada herramienta (ver Tabla 1). Se han enumerado cada una de las categorías temáticas ofrecidas por las distintas herramientas de IAGen: *Copilot*, *DeepSeek* y *ChatGPT*. En la primera columna de la tabla se presentan las categorías dadas por Copilot en el mismo orden con que las mostró, en las columnas 2 y 3 se incluyen las categorías coincidentes de *DeepSeek* y *ChatGPT*. Para no perder la información, se coloca al lado de cada una el orden en que fueron dadas por las distintas herramientas de IAGen

Tabla 1. Categorías temáticas ofrecidas por los modelos de IAGen del grupo de discusión.

Herramienta utilizada para el análisis		
<i>Copilot</i>	<i>DeepSeek</i>	<i>ChatGPT</i>
1. Gestión del comportamiento en el aula		1. Gestión del comportamiento disruptivo
	1. Autoridad y Disciplina en el Aula	2. Autoridad y respeto docente
2. Relación con las familias y percepción social del docente	3. Relación Familia-Escuela	3. Relación con las familias 8. Percepción social del profesorado
3. Colegiación y protección legal del profesorado		4. Riesgo de denuncias y necesidad de respaldo institucional
4. Identidad docente y miedo al aula	2. Rol del Profesor y Identidad Docente	5. Identidad docente y miedo al aula
5. Trabajo en equipo y pertenencia al grupo docente		6. Trabajo colaborativo y pertenencia al grupo de profesorado
6. Inclusión y atención a la diversidad	5. Inclusión y Atención a la Diversidad	
7. Participación del docente en actividades lúdicas	4. Innovación Educativa y Metodologías	7. Metodologías y prácticas educativas
8. Violencia y maltrato en el entorno escolar	7. Casos Críticos: Maltrato y Ética Profesional	
9. Evaluación del e-Diario	6. Evaluación del e-Diario	9. Función del e-diario como espacio reflexivo

Fuente: Elaboración propia.

Como puede apreciarse en la tabla anterior, hay una alta coincidencia. Lo primero que se constata es que para *Copilot* y *ChatGPT* hay 9 categorías temáticas relevantes, y para *DeepSeek* hay 7. En los tres casos, las cantidades son bastante similares, lo cual es destacable si partimos de la idea de que en el *prompt* no se indicó cuántas debían aparecer.

Hay cuatro categorías en las que coinciden las tres herramientas. Estas son las que constan en la Tabla 2. Aunque cada herramienta le da una etiqueta diferente, en definitiva están considerando la misma temática.

Tabla 2. Coincidencias de los tres modelos de IAGen en las categorías temáticas.

<i>Copilot</i>	<i>DeepSeek</i>	<i>ChatGPT</i>
2. Relación con las familias y percepción social del docente	3. Relación Familia-Escuela	3. Relación con las familias 8. Percepción social del profesorado
4. Identidad docente y miedo al aula	2. Rol del Profesor e Identidad Docente	5. Identidad docente y miedo al aula
7. Participación del docente en actividades lúdicas	4. Innovación Educativa y Metodologías	7. Metodologías y prácticas educativas
9. Evaluación del e-Diario	6. Evaluación del e-Diario	9. Función del e-diario como espacio reflexivo

Hay seis categorías temáticas en las que coinciden al menos dos de los modelos, y, como se puede observar en la Tabla 3, estas coincidencias no tienen un patrón definido, ya que *Copilot* converge en 5 propuestas, *ChatGPT* en 4 y *DeepSeek* en 3. Hay que recordar que justamente *DeepSeek* fue el modelo que aportó menos categorías. Por último, no hay ninguna categoría que haya sido propuesta exclusivamente por una sola herramienta.

A modo de ejemplo, en la tabla 4 se muestran las respuestas para las dos primeras categorías descritas en los tres modelos a fin de constatar las coincidencias y diferencias de forma y contenido en la información aportada.

Tabla 3. Coincidencias de al menos dos de los modelos de IAGEN en las categorías temáticas.

<i>Copilot</i>	<i>DeepSeek</i>	<i>ChatGPT</i>
1. Gestión del comportamiento en el aula	1. Autoridad y Disciplina en el Aula	1. Gestión del comportamiento disruptivo 2. Autoridad y respeto docente
3. Colegiación y protección legal del profesorado		4. Riesgo de denuncias y necesidad de respaldo institucional
5. Trabajo en equipo y pertenencia al grupo docente		6. Trabajo colaborativo y pertenencia al grupo de profesorado
6. Inclusión y atención a la diversidad	5. Inclusión y Atención a la Diversidad	
8. Violencia y maltrato en el entorno escolar	7. Casos Críticos: Maltrato y Ética Profesional	

Tabla 4. Ejemplo de la información dada por las herramientas de IAGen para una de las categorías temáticas.

<i>Copilot</i>	1. Gestión del comportamiento en el aula	Gestión del comportamiento en el aula • Frecuencia: Muy alta • Subtemas: • Autoridad docente vs. negociación • Reacciones ante la desobediencia • Consecuencias y comunicación con las familias • Ejemplo: “...hay que decirle: después hablamos tú y yo, antes que decirle ¡que no, que no!, porque esto hay que aclararlo ¿no?” (Informante 2)
<i>DeepSeek</i>	1. Autoridad y Disciplina en el Aula	**Autoridad y Disciplina en el Aula** - **Frecuencia:** Alta. - **Ejemplo:** - **"El tema de un niño que no hace caso... la profesora le dijo que se fuera al final de la cola y el alumno no hacía caso... ¿cuál es el papel del maestro?"** (Informante 7). - Discusión sobre estrategias como negociación, consecuencias posteriores (notas a padres) o imposición inmediata.
<i>ChatGPT</i>	1. Gestión del comportamiento disruptivo 2. Autoridad y respeto	1. Muy alta. Constituye uno de los ejes centrales. <i>“Había un chico de tercero de primaria... la profesora le dijo que se fuera al final de la cola y el alumno no hacía caso... ¿qué podemos hacer?”</i> (Informante 7) 2. Alta. Se debate el significado de autoridad y la diferencia con autoritarismo. <i>“Yo pienso que no es lo mismo tener la autoridad que ser autoritario, la autoridad bien entendida y bien llevada es ideal”</i> (Informante 2).

Como puede apreciarse, la información dada por cada herramienta puede diferir en cuanto al formato de presentación, pero tanto la estructura como los elementos se mantienen. Así, en cada caso se presenta la categoría temática, seguida de la

frecuencia. En los tres se aportan comentarios textuales hallados en el texto identificando al informante, tal y como se solicitó en el *prompt*.

Algunos matices son los que se indican a continuación. En el caso de *Copilot* se incluyen los subtemas abordados en el grupo de discusión, dentro del marco de cada categoría. Por ejemplo, en el caso de la categoría “Gestión del comportamiento en el aula”:

- **Subtemas:**
 - Autoridad docente vs. negociación
 - Reacciones ante la desobediencia
 - Consecuencias y comunicación con las familias

En el caso de *ChatGPT* y *DeepSeek*, hacen una síntesis de lo que aborda la categoría temática. Por ejemplo, respecto a la categoría que se muestra en la tabla anterior:

En el caso de *ChatGPT* dice acerca de “Autoridad y respeto docente”: Se debate el significado de autoridad y la diferencia con autoritarismo.

En el caso de *DeepSeek* dice de “Autoridad y Disciplina en el Aula”: Discusión sobre estrategias como negociación, consecuencias posteriores (notas a padres) o imposición inmediata.

Otro elemento destacable, es la posibilidad de análisis comparado desde los resúmenes globales y por participante que facilitan las herramientas. Estos pueden ser otro elemento que nos acerca al pensamiento del grupo y por tanto que nos da una visión general del relato. A continuación, en la tabla 5 se muestran los tres resúmenes aportados.

En los tres resúmenes se mencionan los temas clave que emergen en el grupo de discusión. De hecho, coinciden en incluir el debate, la reflexión compartida y la riqueza del diálogo como forma de llegar a significados compartidos a través de la plataforma de intercambio.

Concretamente, las tres herramientas destacan, con ligeros matices diferenciados, los mismos aspectos que emergen cuando se establece un diálogo entre pares como el que se relata en el texto. Según *CoPilot*, en el texto se destaca la posibilidad de enfrentarse como grupo a dilemas éticos; *DeepSeek* muestra que se evidencian tensiones entre

ideales educativos y realidades del aula y tensiones entre teoría y práctica; y, en el caso de *ChatGPT*, se muestra la confrontación de experiencias y elaborar colectivamente significados. Por último, en los tres resúmenes aparece, de forma coincidente, la valoración del e-diario como herramienta de intercambio y apoyo, como un catalizador para la reflexión crítica y compartida, aunque se sugiere mejorar su estructura para optimizar el intercambio de ideas.

Tabla 5. Resumen general del Grupo de discusión aportado por los tres modelos de IAGen.

Copilot

Resumen General del Grupo de Discusión

El grupo de discusión abordó múltiples dimensiones de la práctica docente desde la experiencia de estudiantes de magisterio de tres universidades españolas. Se debatieron temas como la gestión del comportamiento en el aula, la autoridad del docente, la relación con las familias, la necesidad de colegiación profesional, el miedo inicial al aula, la inclusión educativa, y la participación del profesorado en actividades lúdicas. También se discutieron casos de violencia y maltrato, generando dilemas éticos sobre el rol del practicante. Finalmente, se valoró positivamente la herramienta e-Diario como espacio de reflexión compartida, proponiendo mejoras para su implementación futura.

DeepSeek

***** **Resumen del Contenido del Grupo de Discusión****

El grupo de discusión abordó temas clave en la formación docente, centrándose en desafíos cotidianos como la disciplina, la autoridad del profesor y la relación con las familias. Se discutieron metodologías innovadoras (como el trabajo por proyectos) y su viabilidad, junto con dilemas éticos (maltrato infantil, inclusión). La herramienta e-Diario fue valorada como un espacio de intercambio y apoyo, aunque se señalaron limitaciones en su estructura (ej. saturación de mensajes). Los informantes destacaron la importancia de la reflexión colectiva para su desarrollo profesional, evidenciando tensiones entre ideales educativos y realidades del aula.

***** **Conclusión****

El análisis revela un diálogo rico entre teoría y práctica, donde los futuros docentes negocian su identidad profesional entre modelos tradicionales y enfoques innovadores. El e-Diario emerge como un catalizador para la reflexión crítica, aunque se sugiere mejorar su estructura para optimizar el intercambio de ideas. La ética y la adaptabilidad aparecen como ejes transversales en su discurso.

ChatGPT

El grupo de discusión evidencia que el **e-diario** actuó como catalizador de reflexión compartida, permitiendo a los futuros maestros confrontar experiencias y elaborar colectivamente significados sobre la práctica docente. Los participantes abordaron principalmente la gestión del comportamiento y la construcción de la autoridad, la compleja relación con las familias y la amenaza de denuncias, así como el temor inicial frente al aula y la definición de la identidad profesional. Asimismo, reflexionaron sobre la importancia del trabajo colaborativo entre docentes, las posibilidades y limitaciones de distintas metodologías, y la percepción social de la profesión. El debate muestra tensiones entre lo ideal y lo real, pero también destaca el valor del intercambio en el e-diario como dispositivo formativo para aprender de la experiencia ajena, validar inquietudes comunes y fortalecer la identidad docente en construcción.

Una vez hecho todo esto, que podría asimilarse con un proceso de triangulación, se tiene una fotografía general de lo aparecido en las transcripciones del grupo de

discusión. Información con la cual se puede continuar el procedimiento de análisis habitual.

Discusión

El objetivo principal de este estudio es explorar metodológicamente la eficacia, fiabilidad e implicaciones éticas al realizar un análisis cualitativo temático de una transcripción de un grupo de discusión, comparando tres herramientas representativas de la IAGen —*Copilot*, *DeepSeek* y *ChatGPT*—. A partir de esta evaluación, se buscó aportar una perspectiva empírica sobre los patrones de riesgo y utilidad que surgen al integrar diversos modelos de IAGen en las metodologías cualitativas y ofrecer orientaciones prácticas para un uso responsable y transparente.

Por lo que respecta a la convergencia en la identificación temática y la utilidad inicial de estas herramientas, la aplicación del mismo *prompt* de consulta a los tres modelos arrojó resultados interesantes, mostrando que las herramientas ofrecieron información similar, aunque con matices. La alta coincidencia en las categorías temáticas identificadas valida la utilidad de la IAGen para obtener una visión general previa del relato o corpus de datos, lo cual puede ser de gran ayuda en el proceso de deconstrucción y reconstrucción de la información obtenida. Esta aproximación inicial automatizada puede resultar muy eficaz para investigadores que enfrentan un volumen importante de datos cualitativos, tal como sugieren Bjelobaba et al. (2025). Conviene recordar que calidad de la respuesta de la IA depende de un *prompt* bien formulado (Sandín Esteban & Sánchez Martí, 2024).

Respecto a la precisión del contenido generado, un aspecto clave de la experiencia fue evaluar la capacidad de las herramientas para cumplir con el requisito de transparencia y precisión en la extracción de citas, un factor diseñado para mitigar el riesgo de fabricación de contenido o *alucinaciones* (Tang et al., 2024). Las tres herramientas cumplieron con el requisito explícito del *prompt* de aportar comentarios textuales hallados en el texto identificando al informante. Este nivel de detalle en la extracción de ejemplos textuales que identifican la fuente (el informante) permite la trazabilidad y

la validación de la veracidad de la información generada por la IAGen, en línea con lo que proponen Penabad-Camacho et al. (2024).

En cuanto a las implicaciones éticas y metodológicas, a pesar de la utilidad demostrada de la IAGen para el QDA —lo que coincide con otros estudios (por ejemplo Bjelobaba et al., 2025; Sandín Esteban & Sánchez Martí, 2024), es imperativo considerar los desafíos éticos planteados, como el sesgo inherente a los datos de entrenamiento y la fabricación de contenido. La IAGen funciona prediciendo texto basándose en patrones estadísticos, lo que implica que sus respuestas carecen de comprensión humana, conciencia y razonamiento abstracto, tal como ya recordaban Cruz Argudo et al. (2024). La naturaleza de este estudio permitió mitigar el riesgo de alucinaciones y sesgo al requerir la extracción directa de citas textuales y el análisis comparativo de los resultados de múltiples modelos. No obstante, el desconocimiento de las bases de datos usadas por los modelos de IAGen refiere la necesidad de cuidar la protección de información confidencial o sensible al introducir datos en el sistema para el análisis (Juca-Maldonado, 2023; Penabad-Camacho et al., 2024; Zangrossi et al., 2024).

Conclusiones

La irrupción de la IAGen en el campo de la investigación académica, especialmente en el QDA, plantea tanto una herramienta útil como desafíos éticos que requieren la atención constante del equipo investigador. En este sentido, las orientaciones que ofrece este estudio no proponen sustituir los métodos clásicos o emergentes de análisis del discurso o de contenido. Más bien, la IAGen se posiciona como una vía de aproximación inicial que puede revelar los grandes temas manifiestos en un texto, independientemente de su longitud. El uso de herramientas de IAGen (como *Copilot*, *DeepSeek* y *ChatGPT*) es efectivo para obtener una visión general previa y automatizada de un volumen importante de datos cualitativos. Así, esta rapidez-efectividad puede ser utilizada como una fase inicial del análisis, permitiendo a los investigadores ahorrar tiempo y recursos.

A pesar de la alta coincidencia en las categorías temáticas obtenidas, la IAGen no tiene la capacidad de comprensión ni de razonamiento abstracto en el sentido humano, ya que funciona basándose en la naturaleza probabilística del lenguaje. Por tanto, es fundamental que el equipo o persona investigadora asuma la responsabilidad de comparar, leer en profundidad, y corregir si es el caso los resultados generados por estas herramientas. El análisis final debe ser completado por los métodos de análisis habituales, pero con una mirada más consciente y deliberada respecto al contenido en sí mismo.

Finalmente, la preocupación sobre los posibles sesgos en las respuestas de la IAGen persiste, debido a que estos provienen del material inicial con que se ha alimentado y enriquecido el modelo. Por ello, se debe tener especial cuidado con la protección de información confidencial o sensible al proporcionar datos para el análisis, ya que existe un desconocimiento de las bases de datos usadas por los modelos de IAGen. Esta información debe ser tratada de manera que no incluya datos personales o que permitan la identificación de los participantes del estudio.

En conclusión, la IAGen es una herramienta poderosa para la fase exploratoria y la sistematización inicial de datos cualitativos, pero su integración exitosa en la investigación académica requiere un marco de responsabilidad humana que garantice la integridad y la ética del proceso.

Financiación

Universitat de Barcelona, Vicerrectorado de política docente (Formación, innovación y evaluación docente). Convocatoria de ayudas a Proyectos de Innovación Docente 2006. Proyecto «E-Diari compartit: una experiència col·laborativa en xarxa en l'especialitat de mestres d'Educació Física de la Facultat de Formació del Professorat». Duración curso académico 2006-2007. Referencia 2006PID-UB/18. IP: Marta Capllonch Bujosa.

Referencias

- Amani, S., White, L., Balart, T., Watson, K., & Shryock, K. (2025). *Applying Generative AI for Thematic Analysis: A Model for Researchers Exploring Qualitative Methods* (SSRN Scholarly Paper No. 5128905). Social Science Research Network. <https://doi.org/10.2139/ssrn.5128905>
- Bjelobaba, S., Waddington, L., Perkins, M., Foltýnek, T., Bhattacharyya, S., & Weber-Wulff, D. (2025). Maintaining research integrity in the age of GenAI: An analysis of ethical challenges and recommendations to researchers. *International Journal for Educational Integrity*, 21(1), 18. <https://doi.org/10.1007/s40979-025-00191-w>
- Cook, D. A., Ginsburg, S., Sawatsky, A. P., Kuper, A., & D'Angelo, J. D. (2025). Artificial Intelligence to Support Qualitative Data Analysis: Promises, Approaches, Pitfalls. *Academic Medicine*, 100(10), 1134-1149. <https://doi.org/10.1097/ACM.00000000000006134>
- Cruz Argudo, F., García Varea, I., Martínez Carrascal, J. A., Ruiz Martínez, P. M., Sánchez Campos, A., & Turró Ribalta, C. (2024). *La inteligencia artificial generativa en la docencia universitaria. Oportunidades, desafíos y recomendaciones*. CRUE Universidades Españolas. https://media.licdn.com/dms/document/media/D561FAQGF2mQcYoaY2A/feed-share-document-pdf-analyzed/0/1710973101353?e=1712188800&v=beta&t=OUZptGKDEAf5jHRy49RA-olc9-nJrl_5yg2HmuWL9CI
- DeepSeek. (2024). *DeepSeek* (Versión R1-0528) [Modelo de lenguaje de gran tamaño]. <https://www.deepseek.com>
- Gray, A. (2024). *ChatGPT «contamination»: Estimating the prevalence of LLMs in the scholarly literature* (No. arXiv:2403.16887). arXiv. <https://doi.org/10.48550/arXiv.2403.16887>
- Juca-Maldonado, F. [Fernando]. (2023). El impacto de la inteligencia artificial en los trabajos académicos y de investigación. *Revista Metropolitana de Ciencias Aplicadas*, 6(51), 290-296.
- Mahyoob, M., Algaraady, J., & Alblwi, A. (2023). A Proposed Framework for Human-like Language Processing of ChatGPT in Academic Writing. *International Journal of Emerging Technologies in Learning*, 18(14), 282-293. Scopus. <https://doi.org/10.3991/ijet.v18i14.41725>

- Microsoft. (2025). *Copilot* (Versión GPT-5) [Modelo de lenguaje de gran tamaño].
<https://copilot.microsoft.com>
- Odri, G.-A., & Yoonb, D. J. Y. (2023). Detecting generative artificial intelligence in scientific articles: Evasion techniques and implications for scientific integrity. *Orthopaedics & Traumatology-Surgery & Research*, 109(8).
<https://doi.org/10.1016/j.otsr.2023.103706>
- OpenAI. (2025). *ChatGPT* (Versión GPT-5) [Modelo de lenguaje de gran tamaño].
<https://chat.openai.com/chat>
- Penabad-Camacho, L., Morera-Castro, M., & Penabad-Camacho, M. A. (2024). Guía para uso y reporte de inteligencia artificial en revistas científico-académicas. *Revista Electrónica Educare*, 28(S), 1-41. <https://doi.org/10.15359/ree.28-S.19830>
- Sandín Esteban, M. P., & Sánchez Martí, A. (2024). A case study in artificial intelligence-assisted qualitative analysis: The use of ChatGPT prompts with UNESCO texts. *Revista Catalana de Pedagogia*, 26, 38-63.
<https://doi.org/10.2436/20.3007.01.209>
- Tang, A., Li, K.-K., Kwok, K. O., Cao, L., Luong, S., & Tam, W. (2024). The importance of transparency: Declaring the use of generative artificial intelligence (AI) in academic writing. *Journal of Nursing Scholarship*, 56(2), 314-318.
<https://doi.org/10.1111/jnu.12938>
- Van Noorden, R., & Perkel, J. M. (2023). AI and science: What 1,600 researchers think. *Nature*, 621(7980), 672-675. <https://doi.org/10.1038/d41586-023-02980-0>
- Wheeler, K. (2025). *How to Use Generative AI to Assist the Analysis of Qualitative Data*. SAGE Publications Ltd. <https://doi.org/10.4135/9781036217471>
- Zangrossi, P., Martini, M., Guerrini, F., De Bonis, P., & Spena, G. (2024). Large language model, AI and scientific research: Why ChatGPT is only the beginning. *JOURNAL OF NEUROSURGICAL SCIENCES*. <https://doi.org/10.23736/S0390-5616.23.06171-4>
- Zohery, M. (2023). ChatGPT in Academic Writing and Publishing: A Comprehensive Guide. En *Artificial Intelligence in Academia, Research and Science: ChatGPT as a Case Study* (pp. 10-61). Novabret Publishing.
<https://doi.org/10.5281/zenodo.7803703>