

# CONCORDANCIA ENTRE PROCESOS DE CODIFICACIÓN CUALITATIVA HUMANA Y CODIFICACIÓN CUALITATIVA AUTOMATIZADA BASADA EN INTELIGENCIA ARTIFICIAL

## AGREEMENT BETWEEN HUMAN QUALITATIVE CODING PROCESSES AND AI-BASED AUTOMATED QUALITATIVE CODING

David ÁLVAREZ-MALDONADO<sup>1</sup>  
Anna MILANO-MENESES<sup>2</sup>

<b>Recibido</b>	: 07/01/2025
<b>Aprobado</b>	: 15/06/2025
<b>Publicado</b>	: 21/06/2025

**RESUMEN:** La inteligencia artificial ofrece mejoras a los métodos tradicionales de investigación cualitativa, en lo que se refiere a la dificultad de trabajar con grandes volúmenes de datos y la fiabilidad de sus resultados. Esta investigación explora el potencial del análisis cualitativo automatizado impulsado por inteligencia artificial, mediante la comparación de dos procesos de codificación paralelos de datos no estructurados compuestos por respuestas textuales abiertas: uno automatizado mediante inteligencia artificial y otro tradicional mediante cognición humana. Se aplicó un cuestionario abierto a una muestra de 263 fanáticos de Disney para comprender sus percepciones sobre lo que representa la marca para ellos mismos, mediante una pregunta de libre respuesta. En el proceso de codificación automatizado se utilizó Python y un modelo de lenguaje denominado Llama 3.2-1b-Instruct. Los resultados mostraron que las codificaciones fueron muy similares en el conjunto de casos, pero de concordancia moderada en lo particular del caso a caso. Se concluye que la inteligencia artificial demuestra potencialidad en la eficiencia del análisis y la escalabilidad, pero evidenció sus límites al exponer inconsistencias en sus resultados, introduciendo redundancias en el proceso de codificación y destacando la necesidad de supervisión mediante procesos cognitivos humanos.

**Palabras clave:** Investigación Cualitativa, codificación, inteligencia artificial, modelos de lenguaje.

**ABSTRACT:** Artificial intelligence offers significant enhancements to traditional qualitative research methods, particularly in handling large volumes of data and improving result reliability. This study explores the potential of AI-driven automated qualitative analysis by comparing two parallel coding processes of unstructured data composed of open-ended textual responses: one automated using artificial intelligence and the other conducted traditionally through human cognition. An open-ended questionnaire was administered to a sample of 263 Disney fans to understand their perceptions of what the brand represents to them, through a free-response question. The automated coding process employed Python and a language model called Llama 3.2-1b-Instruct. The results showed that while the coding outcomes were highly similar across the dataset, there was only moderate agreement at the individual case level. It is concluded that artificial intelligence demonstrates strong potential in

<sup>1</sup> Antropólogo. Universidad Tecnológica Metropolitana, [david.alvarez@utem.cl](mailto:david.alvarez@utem.cl) y ORCID: <https://orcid.org/0000-0002-6375-0461>

<sup>2</sup> Ingeniero Comercial. Instituto de Capacitaciones J21, [contacto@juego21.com](mailto:contacto@juego21.com) y ORCID: <https://orcid.org/0009-0006-1548-9330>

terms of analytical efficiency and scalability, but also reveals limitations by introducing inconsistencies and redundancies in coding, underscoring the need for oversight through human cognitive processes.

**Keywords:** Qualitative Research, coding, artificial intelligence, language models.

## INTRODUCCIÓN

El análisis cualitativo se destaca por la capacidad de abordar fenómenos complejos y datos no estructurados (Jiménez-Partearroyo et al., 2024; Mees-Buss et al., 2022; Magnani y Gioia, 2023), generando una profundidad y detalle en análisis del material empírico que no logran otros enfoques metodológicos (Alvesson y Karreman, 2000; Gioia et al., 2013). Sin embargo, la metodología cualitativa ha tenido críticas válidas y límites históricos reconocidos ampliamente, como la dificultad de trabajar con grandes volúmenes de información y el relativismo subjetivo de los resultados, los cuales limitan la fiabilidad y generalización de las conclusiones (Fileri et al., 2022; Marcolin et al., 2023; Schmitt, 2024). En este contexto, los procesos automatizados de investigación cualitativa, impulsados por inteligencia artificial, emergen como una alternativa para superar tanto los límites para enfrentar grandes volúmenes de información de forma eficiente, como para mejorar la fiabilidad de sus resultados (Christou, 2024; Fileri et al., 2022; Marcolin et al., 2023; Schmitt, 2024).

Para explorar esta situación, se pone a prueba la investigación cualitativa con procesos automatizados basados en inteligencia artificial, mediante la comparación de sus resultados con los derivados de la investigación cualitativa tradicional basada en la cognición humana (Marcolin et al., 2023). Utilizando una muestra de 263 personas que respondieron textualmente un cuestionario abierto que les preguntaba sobre lo que representa la marca Disney para ellos, se realizan dos procesos de codificación paralelos a la diversidad de respuestas de los fanáticos. Estos datos cualitativos no estructurados fueron una prueba tanto para la inteligencia artificial como para el proceso de cognición humana, con el propósito de evidenciar similitudes y diferencias en sus logros.

El alcance de esta investigación es exploratorio, utilizando el índice de concordancia de Kappa de Cohen para verificar qué tan similares o diferentes fueron los resultados de ambos

procesos. De esta manera, se contribuye en la reflexión sobre el rol y el potencial de la inteligencia artificial en la investigación científica de enfoque cualitativo. Primero se aborda una revisión de la literatura especializada en la materia de investigación cualitativa utilizando procesos automatizados impulsados por inteligencia artificial, luego se detalla las especificaciones de la metodología aplicada, posteriormente se abordan los resultados obtenidos de la comparación de los dos procesos de codificación, y finalmente se concluye que ambos procesos son complementarios, siendo útil mantener un enfoque híbrido que combine automatización y prácticas de análisis tradicionales, logrando mejorar la escalabilidad y fiabilidad de los análisis cualitativos.

## **DESARROLLO**

En la actualidad, se reconoce transversalmente que el análisis cualitativo tiene ventajas y fortalezas en la investigación científica (Christou, 2024; Filieri et al., 2022; Gioia, 2021; Jiménez-Partearroyo et al., 2024; Marcolin et al., 2023; Mees-Buss et al., 2022; Magnani y Gioia, 2023; Schmitt, 2024; Lexman et al., 2024), por ejemplo, genera información a partir de datos no estructurados de mayor profundidad y riqueza en comparación con otros enfoques metodológicos, sin embargo, también tiene importantes limitaciones o debilidades, como es el caso de que requiere un intensivo uso de energía y tiempo por parte de los investigadores, siendo procesos de investigación poco eficientes (Filieri et al., 2022), además de las críticas respecto a la subjetividad y el relativismo poco fiable de los resultados (Schmitt, 2024).

En este contexto, las debilidades indicadas del enfoque cualitativo pueden ser abordadas y superadas mediante la inclusión de procesos automatizados basados en inteligencia artificial en las prácticas de análisis, existiendo diferentes investigaciones que demuestran el potencial de escalabilidad y eficiencia para el análisis cualitativo impulsado por procesos automatizados (Christou, 2024; Marcolin et al., 2023; Schmitt, 2024) o la posibilidad de incorporar mayores grados de confiabilidad y validez en los procesos de análisis, disminuyendo el relativismo derivado de la subjetividad del investigador (Schmitt, 2024).

Si bien existen estas oportunidades de mejoras en el análisis cualitativo asistido por procesos automatizados impulsados por inteligencia artificial, es importante destacar que esos potenciales avances no reemplazan los análisis basados en las capacidades cognitivas humanas, sosteniéndose la importancia de mantener el análisis desde un enfoque híbrido que mezcle las operaciones cognitivas humanas de análisis tradicional y las operaciones automáticas impulsadas por la inteligencia artificial (Filieri et al., 2022; Marcolin et al., 2023; Gao et al., 2023).

Las causas de esta necesidad de enfoque híbrido en la incorporación de procesos automatizados en el análisis cualitativo se encuentran en los posibles límites interpretativos de los procesos automatizados basados en inteligencia artificial (Filieri et al., 2022), en comparación con las capacidades interpretativas de las operaciones cognitivas humanas. Además, el análisis automatizado impulsado por inteligencia artificial ha demostrado problemas en reconocer emociones o secuencias de tiempo en el análisis discursivo (Filieri et al., 2022; Marcolin et al., 2023). Junto con esto, el análisis automatizado incorporado en la investigación cualitativa se estima que genera una excesiva homogenización debido a la estandarización, que limita la variabilidad del análisis (Gao et al., 2023), existiendo un mayor potencial de diversidad en los procesos de análisis humanos. En este marco, hay quienes sostienen que la inteligencia artificial no genera análisis de suficiente calidad en comparación con el análisis humano, evidenciándose que la inteligencia artificial no puede replicar las capacidades de la experiencia humana en la investigación cualitativa (Gibson y Beattie, 2024).

Independiente de los diferentes límites identificados, y de las críticas a los procesos automatizados con finalidad de análisis cualitativo, es posible sostener la relevancia de la combinación del análisis cualitativo basado en procesos cognitivos humanos y análisis cualitativo basados en procesos automatizados mediante inteligencia artificial, dado que en uno y otro caso, mejora los resultados y superan sus propios límites en relación a los problemas de eficiencia y de profundidad interpretativa (Christou, 2024; Filieri et al., 2022; Marcolin et al., 2023; Schmitt, 2024). Es necesario considerar que los límites del análisis cualitativo tradicional son fundamentalmente sus escasas posibilidades de escalamiento al

momento de aumentar el tamaño de la muestra y la crítica al relativismo subjetivo del análisis basado en un investigador humano (Marcolin et al., 2023; Schmitt, 2024). El enfoque híbrido que combina procesos cognitivos humanos y procesos automatizados basados en inteligencia artificial permiten abordar estas dos brechas, aumentando el potencial de escalabilidad e indicando posibles sesgos del análisis tradicional (Marcolin et al., 2023). De esta forma, junto con la superación de las pequeñas muestras no generalizables de los estudios cualitativos, la inteligencia artificial puede proporcionar una validación a la naturaleza subjetiva y relativa de la investigación cualitativa, superando la intuición humana (Schmitt, 2024), a la vez que la capacidad humana puede supervisar y mejorar los resultados del proceso automatizado.

En este contexto, hay quienes sostienen que la investigación cualitativa tradicional podría quedar obsoleta, siendo reemplazada por una ingeniería de datos cualitativos no estructurados (Schmitt, 2024), sin embargo, también hay quienes plantean que la inteligencia artificial tiene límites que deben ser abordados mediante un enfoque híbrido (Christou, 2024; Filieri et al., 2022; Marcolin et al., 2023), dado que el enfoque crítico del análisis tradicional basado en las capacidades cognitivas del humano sigue siendo irremplazable, permitiendo potenciar el análisis humano en vez de reemplazarlo (Christou, 2023a; Christou, 2023b; Christou, 2024). El preliminar consenso respecto de este debate es que es necesario que se supere el escepticismo y la abstinencia del uso de los procesos automatizados en inteligencia artificial, siendo necesario combinar los métodos tradicionales de investigación cualitativa con las nuevas tecnologías impulsadas por la inteligencia artificial (Schmitt, 2024).

En este sentido, el escepticismo señalado se ha reducido rápidamente, lo cual se evidencia en que la inteligencia artificial se ha utilizado ampliamente y de forma creciente en la investigación científica (Christou, 2023a; Christou, 2023b; Gebreegziabher et al., 2023; Jeldes-Delgado et al., 2024), por ejemplo, con fines predictivos (Kumbure et al., 2022) o con fines formativos (Palea et al., 2024; Sinha et al., 2024). Además, se ha destacado el potencial del trabajo colaborativo en los procesos de codificación, lo cual puede mejorar la concordancia más eficaz y eficiente de la codificación, en menos tiempo y con mayor porcentaje de acuerdo a Gao et al. (2023).

En estos procesos de codificación híbridos, mezclando operaciones cognitivas humanas y operaciones automatizadas basadas en inteligencia artificial, muestran que los resultados presentan similitudes y diferencias entre el análisis humano y el generado mediante inteligencia artificial, sucediendo que la codificación humana reconoce algunos contenidos que no reconoce la inteligencia artificial, y viceversa, la inteligencia artificial reconoce contenidos que no reconoce la operación cognitiva humana (Hamilton et al., 2023). En este contexto, la presente investigación busca contribuir en esta temática, teniendo como objetivo comparar el análisis humano y de la inteligencia artificial, con el propósito de identificar sus similitudes y diferencias, contribuyendo a explorar este ámbito operativo en el análisis de datos cualitativos.

## **METODOLOGÍA**

Esta investigación tiene el objetivo de explorar el potencial del análisis cualitativo automatizado impulsado por inteligencia artificial, mediante la comparación entre el análisis cualitativo humano y el análisis cualitativo de la inteligencia artificial (Marcolin et al., 2023), basándose en la identificación de concordancias en los procesos de codificación (Kull, 2020). Con este propósito, se aplica un cuestionario abierto a personas que son parte de la audiencia de las películas Disney, identificándolos en comunidades virtuales de fanáticos de esta marca de películas, en el que se contactaron informantes clave definidos como personas influyentes en redes sociales, las cuales estaban asociadas a Disney, por ejemplo, creadores de contenido sobre Disney, tatuadores de iconografía Disney, líderes de opinión de comunidades de fanáticos de Disney, entre otros tipos de informantes. Para abordar este levantamiento de información, se utiliza un enfoque cualitativo de tipo netnográfico (Kozinets, 1998; Kozinets, 2002; Kozinets, 2006; Kozinets et al., 2018), el cual es una metodología validada de trabajo en terreno en entornos virtuales para el acceso de datos cualitativos. Estas personas contactadas tomaron el rol de informantes clave del proceso investigativo, y se les solicitó la colaboración de publicar un cuestionario online realizado en Google Forms en sus redes sociales, con el propósito de que sus seguidores y cercanos, partes de la comunidad virtual de fanáticos Disney, contestaran un cuestionario abierto con la pregunta: ¿Qué representa

para ti la marca Disney? En este contexto, los participantes del cuestionario pudieron expresarse de forma escrita libremente, desarrollando sus ideas sobre lo que representa para ellos la marca Disney, recopilando de esta forma datos cualitativos no estructurados para realizar esta investigación, los cuales requieren de procesos de codificación.

De esta forma, las respuestas se constituyeron en datos cualitativos no estructurados sobre la percepción de los fanáticos de Disney respecto a lo que representa la marca de películas y series para ellos mismos, desde su propio criterio subjetivo. En este sentido, es importante destacar que los datos cualitativos generados se recolectaron de forma abierta y no estructurada, con una respuesta única por cada participante, la cual pudo contener una extensión variable de contenidos, según el criterio del participante, sin generar un límite reducido de palabras. La muestra final se compuso de 263 personas, de las cuales el 89,4% se identificaron con el género femenino, mientras que el 10,6% se identificaron con el género masculino. La diferencia del género en los voluntarios no fue objeto de análisis en esta investigación, pero permite suponer que la comunidad del caso de estudio es mayormente femenina, o que los voluntarios autoseleccionados tienen un perfil mayormente femenino. Los rangos etarios declarados fueron los siguientes: 1,1% de menores de 14 años; 6,1% de personas entre 15 a 17 años; 28,5% de 18 a 24 años; 51% de 25 a 34 años; 12,5% de 35 a 44 años; y 0,8% mayores de 45 años. De esta forma, también es posible concluir que el mayor porcentaje de personas interesadas en contestar este cuestionario voluntario son jóvenes o adultas jóvenes. Finalmente, el consentimiento informado se aplicó en el mismo instrumento.

El procedimiento de codificación humano fue inductivo y emergente, analizado las respuestas escritas de los participantes, para asignarlos a una categoría representativa del conjunto de significados de la respuesta. Las codificaciones similares se agruparon en categorías representativas de temáticas comunes en las respuestas. El propósito de la codificación es estructurar la información mediante patrones de significados comunes que agrupan diferentes casos. El procedimiento de codificación automatizado basado en inteligencia artificial también fue inductivo y emergente, sin considerar ninguna categorización previa o instrucción preliminar sobre los códigos aplicables en el análisis. Se utilizó el lenguaje programación Python (Bird et al., 2009; Chollet, 2021) mediante el

software Synder y el modelo de inteligencia artificial llama-3.2-1b-instruct mediante LM Studio. El prompt utilizado es el siguiente: Eres un asistente de *text analytics* que usará sus habilidades de entendimiento profundo del lenguaje. Tu tarea será revisar comentarios de fanáticos de Disney sobre lo que representa la marca Disney para ellos mismos, y encontrar los tópicos que se mencionen. Un comentario puede contener uno o más tópicos, por lo que, por favor, necesito que identifiques exclusivamente el más importantes, identificando un tópico general por comentario.

Finalmente, se agrupan las diferentes codificaciones generadas según temáticas o tópicos, en función de los elementos comunes, estructurando los datos cualitativos generados por el cuestionario mediante la identificación de patrones de significados comunes. Finalmente se aplica la prueba de Chi-Cuadrado y la prueba de Kappa de Cohen para verificar la concordancia entre la codificación humana y la de la inteligencia artificial, concluyendo en una interpretación de las similitudes y diferencias de los resultados de ambos procesos de codificación.

## RESULTADOS

La codificación cualitativa tradicional basada en procesos cognitivos humanos logró identificar seis temáticas generales que agrupaban las diferentes respuestas de los participantes, los cuales son las siguientes:

1) Nostalgia y recuerdos infantiles, asociados a que los participantes mencionan que Disney les evoca recuerdos nostálgicos y felices de su infancia. Este tema se comprende bajo la lógica de que los productos Disney, sus historias y personajes, ejerce un estímulo que genera recuerdos asociados a experiencias pasadas de la infancia en que estos productos fueron relevantes. Algunos ejemplos de segmentos de respuestas de los participantes del cuestionario son los siguientes: “Conexión con el niño interno” (Participante 1); “Disney significa mi infancia y muy buenos recuerdos, mi cercanía al cine que es a lo que me quiero dedicar ahora por lo tanto mi pensamiento y sentimientos son de nostalgia ya que cada vez que veo una película antigua o nueva me vuelvo a sentir la niña que fui” (Participante 5); “Amo con mi vida Disney porque evoca una profunda nostalgia y me transporta a momentos

felices de mi infancia” (Participante 49); “Disney es mi infancia y esa niña interior que siempre llevaré” (Participante 96); y, “Disney para mí representa volver a reencontrarte con tu niña/o interior, es una sensación de nostalgia al recordar los buenos tiempos creciendo junto a la marca” (Participante 185).

2) Magia y fantasía, situación en que Disney es visto como un universo fantástico y un escape de la rutina cotidiana, donde la fantasía puede desarrollarse. Algunos ejemplos de este tema son las siguientes frases: “Representa mi lugar de escape de la realidad. Actualmente, me sigue gustando escaparme de la realidad adulta y disfrutar de las películas e historias” (Participante 40); “La magia y los sueños si se pueden hacer realidad! ¡Pude cumplir el sueño de viajar a Orlando y fue lo más mágico y maravilloso del mundo! Aun no lo supero, creo que para alguien que es fans es muy difícil superarlo jaja, entendí que importante es creer y soñar, así que el nunca dejes de soñar se me quedo grabado forever” (Participante 66); “Disney para mí representa magia y felicidad, es sentir y ver que todo es posible, en cada película que tiene, puedes encontrar cosas maravillosas” (Participante 124); “Disney representa un sueño, es como lograr lo inalcanzable, felicidad y emoción” (Participante 156); y, “Representa la capacidad de escapar un tiempo de la realidad mediante sus historias de fantasía” (Participante 183).

3) Emociones positivas, en que Disney es asociado a la generación de alegría y otros sentimientos positivos, como tranquilidad, representando un lugar seguro y positivo emocionalmente. A continuación, se mencionan ejemplos de segmentos relevantes de esta temática: “Paz, me da mucha tranquilidad. Es muy lugar seguro” (Participante 52); “Sentimientos de alegría, felicidad, emoción, cosas bonitas” (Participante 109); “Disney para mí es felicidad, siempre que pienso en algo relacionado con ello soy una persona muy feliz” (Participante 176); y, “Me tranquiliza saber que la mayoría de las historias tienen finales felices” (Participante 249).

4) Enseñanza y valores, en que los participantes valoran las lecciones de vida y los valores que transmiten los productos de Disney, es decir, sus historias y personajes. En este sentido, los productos Disney generan un impacto en las creencias y comportamientos de los consumidores o fanáticos. Esta temática se puede representar en las siguientes frases:

“Representa lo que soy hoy en día, sus emociones y enseñanzas de cada película tanto Disney como Pixar me llegan y algunas hasta me emocionan al punto de llorar. Hacen que reflexione de mi día a día o el cómo sobrellevar algunos problemas de mi vida” (Participante 57); “Disney representa valores como los que aparecen en el Rey León y Mulán” (Participante 77); “Que te enseña cosas maravillosas, lecciones que tenemos que aprender” (Participante 136); y, “Es un mundo del cual no soy muy fanático, pero hay ciertas películas que me dejaron algunas enseñanzas en mi vida” (Participante 255).

5) Conexión con la familia, en que Disney representa la unión parental y los recuerdos compartidos con los seres queridos de diferentes generaciones, estableciéndose como un mediador y un puente emocional entre los diferentes auditores. Este contenido se puede evidenciar en las siguientes frases ejemplos: “Disfrutar en familia” (Participante 90); “Me hace pensar en las veces que, en familia, mis papás nos acompañaban a ver las pelis en VHS. Disney es algo con lo que crecí. Sus historias, sobre todo” (Participante 105); y, “Familia, tiempo de calidad y risas” (Participante 201).

6) Un estilo de vida personal, en que Disney representa más que sólo entretenimiento, siendo un estilo de vida y un elemento constitutivo de la identidad. Esto se puede observar en que algunos participantes hacen referencia a coleccionar productos o estructuración de la vida cotidiana inspirados en lo que perciben de Disney, lo que significa que la marca se integra fuertemente en los consumidores. Este tema se puede observar en las siguientes frases relevantes: “Representa todo, me encanta coleccionar cosas de Disney” (Participante 14); “Es parte de quien soy” (Participante 20); “Representa un estilo de vida, una manera de vivir, mi forma de pensar. Mis referencias en varios ámbitos de la vida” (Participante 21); y “Disney es lo que me representa, en el sentido que, si tienen que asociar algo a mí, es Disney” (Participante 22)

En este marco, se codificaron las seis diferentes categorías mediante un número indicativo del 1 al 6. A continuación, en la Tabla 1 se indica la frecuencia y porcentaje de cada codificación en el proceso basado en el humano:

**Tabla 1***Codificación tradicional de representaciones de la marca Disney*

<b>Codificación tradicional de Respuestas a la pregunta</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
<b>¿Qué representa para ti la marca Disney?</b>		
1) Nostalgia y recuerdos infantiles	128	48,7
2) Magia y fantasía	41	15,6
3) Emociones positivas	59	22,4
4) Enseñanza y valores	8	3,0
5) Conexión con la familia	10	3,8
6) Estilo de vida personal	17	6,5
Total	263	100,0

El proceso de codificación automatizado basado en inteligencia artificial no generó una homogenización de las codificaciones ni redujo la variabilidad, como se advertía en la literatura (Gao et al., 2023), todo lo contrario, generó más categorías de codificación que las necesarias, generando traslapes con falta de discriminación y redundancias, por ejemplo: estableció el tópico de recuerdos de la infancia, por separado del tópico de nostalgia de la infancia, los cuales pueden ser temáticas comunes, aunque uno tenga sentimientos de nostalgia, y otro haga referencia al proceso cognitivo de recordar, dado que la nostalgia involucra necesariamente el proceso cognitivo del recuerdo, es decir, los recuerdos son un conjunto que involucra el conjunto de la nostalgia, generándose la necesidad de codificarlos con la misma categoría debido al traslape de significados; también generó diferencias innecesarias en relación a mencionar tópicos sobre recuerdos de la familia, de forma diferenciada con conexión con la familia, entre otros códigos con variabilidad innecesaria y redundancias, en que no se establecieron adecuadamente las generalizaciones de patrones comunes de significados.

Una posible explicación de estos traslapes con falta de discriminación se puede encontrar en la literalidad del proceso automatizado de codificación, dado que se configuraron parámetros del modelo que minimizaba su creatividad, reduciendo el nivel de aleatoriedad en las respuestas generadas, estableciendo un parámetro de temperatura de inteligencia artificial

con valores bajos cercanos al cero, generando que el modelo sea más determinista, eligiendo las palabras más probables según el entrenamiento. La temperatura es un parámetro de configuración de la inteligencia artificial y el modelo de lenguaje utilizado. Se estableció una temperatura configurada al 0,001 la cual demostró ser bastante literal y poco creativa. Debía ser un número más equilibrado, por ejemplo 0,5 o similar. Esta opción habría permitido generar respuestas más creativas y menos literales. Probablemente el trabajo de codificación requiera que sean procesos creativos más que deterministas técnicamente, los cuales serían muy apegados a la literalidad de la palabra, impidiendo la generalización mediante interpretación de significados y patrones. Este asunto explica la ventaja de los procesos de cognición humana en estas tareas versus los procesos automatizados antes de la inteligencia artificial, dado que somos capaces de trabajar con diferentes niveles de temperatura, de forma flexible, sin requerir configuración.

Junto con esto, también se encontraron errores de codificación, generando codificaciones aisladas con una sola respuesta. Por ejemplo, el participante 14 que indicó “Representa todo, me encanta coleccionar cosas de Disney”, fue codificado bajo la categoría de coleccionista, lo cual, al ser una codificación con un solo caso, pierde el propósito técnico de la codificación de agrupar diferentes elementos y respuestas comunes en un significado más amplio. En el caso de la codificación humana, el caso 14 fue interpretado bajo el código de estilo de vida y cultura personal. Este tipo de errores fue poco frecuente, alcanzando escasos 9 casos, es decir, un 3,4% de la muestra, los cuales fueron recodificados con el número cero, para diferenciarlos de las otras codificaciones.

Finalmente, la codificación automatizada también generó aproximadamente seis categorías de codificación, pero requirió una asistencia crítica de procesamiento cognitivo humano que reagrupara las categorías redundantes, por ejemplo, uniendo nostalgia por la infancia con recuerdos de la infancia. Sin la asistencia humana, la alta variabilidad con poca discriminación no sería útil para el análisis cualitativo, estableciéndose finalmente que el resultado generado requiere de supervisión. Probablemente si se le hubieran indicado las seis categorías de forma previa, como un input para que clasificara las respuestas deductivamente, en vez de emerger inductivamente estas categorías, se habría solucionado este problema. En

este sentido, es posible que sea necesario aplicar un proceso preliminar de codificación inductiva humana para entrenar un procesamiento automatizado deductivo, en vez de permitirle realizar un proceso inductivo y emergente similar al humano. Para complementar ambos procesos, el de codificación tradicional, con el de codificación humana, se podría primero generar una etapa de codificación humana inductiva y emergente, para posteriormente escalar a una muestra de mayor tamaño mediante codificación automatizada de forma deductiva y entrenada con la primera codificación.

Es importante destacar que, a pesar de los errores de codificación generados por la inteligencia artificial, los resultados del proceso de codificación automatizada fueron muy similares al proceso de codificación humana, exceptuando los errores de codificaciones de un 3,4% y el hecho de que no siempre se codificó bajo el mismo contenido que el proceso humano, generando similares categorías, pero diferentes aplicaciones de estas al momento de codificar. Esto se debe a que las frases de los participantes e interpretaciones de estas pueden ser ambiguas, por ejemplo, tener nostalgia por la infancia también puede ser interpretado como un momento de conexión con la familia, si se considera que la infancia sucede en contexto familiar generalmente, existiendo una ambigüedad en frases respecto a si se consideran de una u otra categoría al momento de la codificación. Probablemente una cantidad relevante de respuestas se ajustan a más de un código. De esta forma, si bien existe una similitud en la generalidad de los datos al momento de codificar, la aplicación caso a caso de estos códigos puede diferir debido a traslapes de patrones de significados comunes.

A continuación, en la Tabla 2 se indican las frecuencias y porcentajes de la codificación automatizada impulsada por la inteligencia artificial, los cuales reportaron los mismos códigos. Los errores del proceso de codificación fueron codificados con el número cero.

**Tabla 2**

*Codificación automatizada de representaciones de la marca Disney*

<b>Codificación automatizada de Respuestas a la pregunta ¿Qué representa para ti la marca Disney?</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
0. Errores de codificación no agrupables	9	3,4
1. Nostalgia y recuerdos infantiles	112	42,6
2. Magia y fantasía	33	12,5

3. Emociones positivas	85	32,3
4. Enseñanza y valores	7	2,7
5. Conexión con la familia	10	3,8
6. Estilo de vida personal	7	2,7
Total	263	100,0

En este contexto, es posible concluir que existen similares codificaciones entre el proceso humano y el proceso automatizado, encontrándose seis códigos en ambos procesos, pero requiriendo una supervisión y limpieza en el resultado del proceso automatizado. Esta similitud se puede verificar mediante el análisis del Chi-Cuadrado de Pearson, que evalúa si existe una asociación estadísticamente significativa entre los datos generados por el proceso humano y los datos generados por el proceso automatizado. El valor de Chi-Cuadrado fue 324.875 con 30 grados de libertad, y p-value de 0,000. Esto indica que la asociación es altamente significativa, sugiriendo que ambos procesos de codificación están asociados. A continuación, en la Tabla 3 se muestra una comparación de las frecuencias y porcentajes de codificación y en la Tabla 4 los resultados del Chi-Cuadrado.

**Tabla 3**

*Comparación de codificación tradicional y automatizada*

Códigos	Frecuencia Inteligencia Artificial	Porcentaje Inteligencia Artificial	Frecuencia Humano	Porcentaje Humano
0	9	3,4	0	0
1	112	42,6	128	48,7
2	33	12,5	41	15,6
3	85	32,3	59	22,4
4	7	2,7	8	3,0
5	10	3,8	10	3,8
6	7	2,7	17	6,5
Total	263	100,0	263	100,0

*Nota.* Los códigos fueron representados con un número del 1 al 6 como variable categórica, donde: 1 representa nostalgia y recuerdos infantiles; 2 representa magia y fantasía; 3 representa emociones positivas; 4 representa enseñanza y valores; 5 representa conexión con la familia; 6 representa un estilo de vida personal; y 0 representa los errores no agrupables del proceso de codificación.

**Tabla 4***Prueba de Chi-Cuadrado de Pearson para la codificación tradicional y automatizada*

<b>Prueba de chi-cuadrado</b>			
	<b>Valor</b>	<b>df</b>	<b>Significación asintótica (bilateral)</b>
Chi-cuadrado de Pearson	324,875 <sup>a</sup>	30	,000
Razón de verosimilitud	242,561	30	,000
Asociación lineal por lineal	62,749	1	,000
N de casos válidos	263		

a. 31 casillas (73,8%) han esperado un recuento menor que 5. El recuento mínimo esperado es ,21.

Respecto de la concordancia entre la codificación humana y la codificación automatizada, se utiliza el índice de Kappa de Cohen, para medir el acuerdo entre ambas evaluaciones, más allá de lo esperado por el azar. El valor alcanzado de Kappa es 0,514 con un p-value de 0,000 indicando un nivel de acuerdo moderado, pero significativo. Cuando dos humanos hacen análisis de codificación cualitativa de forma paralela, las diferencias se deben resolver hasta alcanzar valor de 0,8 de Kappa, por lo que este acuerdo moderado indica una nueva necesidad de supervisión para dirimir las diferencias de codificación. A continuación, en la Tabla 5 se muestran los resultados del índice de Kappa de Cohen:

**Tabla 5***Medida de acuerdo del índice de Kappa de Cohen para el proceso tradicional y automatizado*

<b>Medidas simétricas</b>					
		<b>Valor</b>	<b>Error estándar asintótico<sup>a</sup></b>	<b>T aproximada<sup>b</sup></b>	<b>Significación aproximada</b>
Medida de acuerdo	Kappa	,514	,038	14,497	,000
N de casos válidos		263			

a. No se presupone la hipótesis nula.

b. Utilización del error estándar asintótico que presupone la hipótesis nula.

## DISCUSIÓN DE RESULTADOS

Esta investigación realizó una comparación de los procesos de codificación tradicional basada en procesos cognitivos humanos y la codificación automatizada impulsada por la inteligencia artificial (Marcolin et al., 2023), utilizando como caso de estudio las percepciones de los fanáticos de Disney respecto de lo que representa la marca para ellos mismos. En este contexto, ambos procesos de codificación identificaron temáticas muy similares en general, sin embargo, se evidenció un nivel de concordancia moderado al momento de utilizar estas temáticas para codificar las respuestas de los participantes del estudio, en el nivel de lo particular del caso a caso, logrando un índice significativo de Kappa de 0,514. Esto representaría concordancias moderadas, pero significativas, en la interpretación de los datos. Las diferencias entre ambas codificaciones particulares pueden ser explicadas por los traslapes de significado, en que una respuesta puede hacer referencia a dos o más códigos temáticos.

Esta situación de alta similitud general permite sostener que los procesos automatizados de análisis cualitativo tienen un potencial relevante para complementar los procesos tradicionales de análisis cualitativo, a razón de similares capacidades para reconocer temáticas en las respuestas a cuestionarios abiertos (Christou, 2024; Marcolin et al., 2023; Schmitt, 2024). La contribución del proceso automatizado radica en que reduce significativamente el tiempo de trabajo requerido para grandes volúmenes de datos, abordando miles o millones de respuestas en mucho menos tiempo y con menos recursos humanos, logrando un resultado similar a los procesos tradicionales de codificación en cuanto a identificación de temáticas se refiere. Sin embargo, el análisis automatizado generó codificaciones con menor discriminación, resultando en redundancias, incapaz de agrupar significados con sutiles diferencias y similitudes, debido a la literalidad de la interpretación del modelo de lenguaje y la ambigüedad de los datos cualitativos. Esto sugiere la importancia de que el proceso de análisis automatizado sea supervisado por el ser humano. Además, es necesario considerar que el modelo de lenguaje utilizado estaba cuantizado, es decir,

reducido para un rendimiento más eficiente, y que podrían existir modelos con un mejor desempeño, pero que requieren recursos informáticos mayores.

De esta forma, el enfoque híbrido que combina las capacidades cognitivas humanas con herramientas automatizadas es respaldado por los resultados, descartando la idea de que la automatización reemplazaría la actividad de codificación humana (Schmitt, 2024). Mediante esta perspectiva, se pueden reducir las limitaciones de ambos métodos por separado, mejorando la fiabilidad y escalabilidad del análisis cualitativo, gracias a su mayor eficiencia. Sin embargo, actualmente no es posible reemplazar las capacidades críticas y analíticas del proceso tradicional basado en capacidades cognitivas de personas, en lo que su función de supervisión refiere (Gibson y Beattie, 2024).

Una mejora de los límites del proceso automatizado es entrenar al modelo de lenguaje con un proceso de codificación humano de forma previa para que, al analizar grandes volúmenes de datos, tenga orientaciones deductivas al momento de trabajar, en vez de inductivas como fue en el caso de esta investigación. Sin embargo, esto limitaría la emergencia de patrones no intuitivos detectados por la inteligencia artificial. Por esta razón, es relevante generar mayores exploraciones sobre este enfoque híbrido para establecer un estándar científico que permita la replicabilidad, logrando cumplir con las expectativas de mayor escalabilidad y fiabilidad, lo cual permitiría superar barreras históricas del análisis cualitativo relacionadas con el sesgo subjetivo y la dificultad de trabajar con grandes volúmenes de datos.

## CONCLUSIONES

Este estudio comparó la codificación cualitativa tradicional, basada en procesos cognitivos humanos, con la codificación automatizada mediante inteligencia artificial, utilizando como caso de análisis las percepciones de fanáticos de Disney sobre la marca. Los resultados muestran una alta similitud general en la identificación de temáticas, aunque con una concordancia moderada a nivel caso a caso ( $Kappa = 0,514$ ). Esto evidencia que los modelos de lenguaje pueden reconocer patrones temáticos de manera efectiva, aunque con menor capacidad discriminativa en respuestas ambiguas o complejas.

La automatización ofrece ventajas claras en términos de eficiencia y escalabilidad, especialmente útil para grandes volúmenes de datos. Sin embargo, las limitaciones observadas —como la literalidad en la interpretación y la generación de redundancias— refuerzan la necesidad de supervisión humana para garantizar la validez del análisis. Los hallazgos respaldan un enfoque híbrido, donde la inteligencia artificial complementa, pero no sustituye, las capacidades analíticas humanas.

Finalmente, se propone avanzar hacia modelos entrenados previamente con codificación humana para mejorar su desempeño en tareas específicas, sin perder la capacidad inductiva que permite descubrir patrones emergentes. Esto destaca la importancia de establecer estándares científicos que aseguren la replicabilidad, fiabilidad y escalabilidad del análisis cualitativo en la era de la inteligencia artificial.

## REFERENCIAS BIBLIOGRÁFICAS

- Alvesson, M. y Kärreman, D. (2000). Varieties of discourse: On the study of organizations through discourse analysis. *Human relations*, 53(9), 1125-1149. <https://doi.org/10.1177/0018726700539002>
- Bird, S., Klein, E. y Loper, E. (2009). *Natural language processing with Python: analyzing text with the natural language toolkit*. O'Reilly Media, Inc.
- Chollet, F. (2021). *Deep learning with Python*. Simon and Schuster.
- Christou, P. (2023a). A critical perspective over whether and how to acknowledge the use of artificial intelligence (AI) in qualitative studies. *The Qualitative Report*, 28(7), 1981-1991. <https://doi.org/10.46743/2160-3715/2023.6407>
- Christou, P. (2023b). How to use artificial intelligence (AI) as a resource, methodological and analysis tool in qualitative research? *Qualitative Report*, 28(7), 1968-1980. <https://doi.org/10.46743/2160-3715/2023.6406>
- Christou, P. A. (2024). Thematic Analysis through Artificial Intelligence (AI). *The Qualitative Report*, 29(2), 560-576. <https://doi.org/10.46743/2160-3715/2024.7046>
- Filieri, R., Lin, Z., Li, Y., Lu, X. y Yang, X. (2022). Customer emotions in service robot encounters: A hybrid machine-human intelligence approach. *Journal of Service Research*, 25(4), 614-629. <https://doi.org/10.1177/10946705221103937>

- Gao, J., Choo, K. T. W., Cao, J., Lee, R. K. W. y Perrault, S. (2023). CoAICoder: Examining the effectiveness of AI-assisted human-to-human collaboration in qualitative analysis. *ACM Transactions on Computer-Human Interaction*, 31(1), 1-38. <https://doi.org/10.1145/3617362>
- Gebreegziabher, S. A., Zhang, Z., Tang, X., Meng, Y., Glassman, E. L. y Li, T. J. J. (2023). Patat: Human-ai collaborative qualitative coding with explainable interactive rule synthesis. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (pp. 1-19). <https://doi.org/10.1145/3544548.3581352>
- Gibson, A. F. y Beattie, A. (2024). More or less than human? Evaluating the role of AI-as-participant in online qualitative research. *Qualitative Research in Psychology*, 21(2), 175-199. <https://doi.org/10.1080/14780887.2024.2311427>
- Gioia, D. (2021). A systematic methodology for doing qualitative research. *The Journal of Applied Behavioral Science*, 57(1), 20-29. <https://doi.org/10.1177/0021886320982715>
- Gioia, D. A., Corley, K. G. y Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research: Notes on the Gioia methodology. *Organizational research methods*, 16(1), 15-31. <https://doi.org/10.1177/1094428112452151>
- Hamilton, L., Elliott, D., Quick, A., Smith, S. y Choplin, V. (2023). Exploring the use of AI in qualitative analysis: A comparative study of guaranteed income data. *International Journal of Qualitative Methods*, 22. <https://doi.org/10.1177/16094069231201504>
- Jeldes-Delgado, F., Ferreira, T. A., Diaz, D. y Ortiz, R. (2024). Exploring gender stereotypes in financial reporting: An aspect-level sentiment analysis using big data and deep learning. *Heliyon*, 10(20). <https://doi.org/10.1016/j.heliyon.2024.e38915>
- Jiménez-Partearroyo, M., Medina-López, A. y Rana, S. (2024). Business intelligence and business analytics in tourism: insights through Gioia methodology. *International Entrepreneurship and Management Journal*, 20, 2287-2321. <https://doi.org/10.1007/s11365-024-00973-7>
- Kozinets, R. V. (1998). On netnography: Initial reflections on consumer research investigations of cyberculture. *Advances in consumer research*, 25(1), 366-371.
- Kozinets, R. V. (2002). The field behind the screen: Using netnography for marketing research in online communities. *Journal of marketing research*, 39(1), 61-72. <https://doi.org/10.1509/jmkr.39.1.61.18935>
- Kozinets, R. V. (2006). Click to connect: Netnography and tribal advertising. *Journal of advertising research*, 46(3), 279-288. <https://doi.org/10.2501/S0021849906060338>
- Kozinets, R. V., Scaraboto, D. y Parmentier, M. A. (2018). Evolving netnography: How brand auto-netnography, a netnographic sensibility, and more-than-human netnography can transform

- your research. *Journal of Marketing Management*, 34(3-4), 231-242.  
<https://doi.org/10.1080/0267257X.2018.1446488>
- Kull, K. (2020). Codes: Necessary, but not sufficient for meaning-making. *Constructivist Foundations*, 15(2), 137-139.
- Kumbure, M. M., Lohrmann, C., Luukka, P. y Porras, J. (2022). Machine learning techniques and data for stock market forecasting: A literature review. *Expert Systems with Applications*, 197.  
<https://doi.org/10.1016/j.eswa.2022.116659>
- Lexman, R. R., Krishnan, G., Baral, R. y Thomas, S. C. (2024). Navigating the digital classroom: a qualitative content analysis of MOOC discourses in Indian e-newspapers. *Journal of Information, Communication and Ethics in Society*, 22(4), 494-516.  
<https://doi.org/10.1108/JICES-03-2023-0027>
- Magnani, G. y Gioia, D. (2023). Using the Gioia Methodology in international business and entrepreneurship research. *International Business Review*, 32(2).  
<https://doi.org/10.1016/j.ibusrev.2022.102097>
- Marcolin, C. B., Diniz, E. H., Becker, J. L. y de Oliveira, H. P. G. (2023). Who knows it better? Reassessing human qualitative analysis with text mining. *Qualitative Research in Organizations and Management: An International Journal*, 18(2), 181-198.  
<https://doi.org/10.1108/QROM-07-2021-2173>
- Mees-Buss, J., Welch, C. y Piekkari, R. (2022). From templates to heuristics: How and why to move beyond the Gioia methodology. *Organizational Research Methods*, 25(2), 405-429.  
<https://doi.org/10.1177/1094428120967716>
- Palea, D., Vadhul, G. y Lee, D. T. (2024). Annota: Peer-based AI Hints Towards Learning Qualitative Coding at Scale. En *Proceedings of the 29th International Conference on Intelligent User Interfaces* (pp. 455-470). <https://doi.org/10.1145/3640543.3645168>
- Schmitt, B. (2024). Transforming qualitative research in phygital settings: the role of generative AI. *Qualitative Market Research: An International Journal*, 27(3), 523-526.  
<https://doi.org/10.1108/QMR-08-2023-0107>
- Sinha, R., Solola, I., Nguyen, H., Swanson, H. y Lawrence, L. (2024). The Role of Generative AI in Qualitative Research: GPT-4's Contributions to a Grounded Theory Analysis. En *Proceedings of the Symposium on Learning, Design and Technology* (pp. 17-25).  
<https://doi.org/10.1145/3663433.3663456>