OBJN
Online Brazilian Journal of Nursing
ISSN: 1676 - 4285 | www.objnursing.uff.br

Envío: 03-Sep-2026
Aprobado: 20-Oct-2026

AGENTES DE INTELIGENCIA ARTIFICIAL PARA EL ANÁLISIS TEMÁTICO REFLEXIVO SIN PROGRAMACIÓN

AGENTES DE INTELIGÊNCIA ARTIFICIAL PARA ANÁLISE TEMÁTICA REFLEXIVA SEM PROGRAMAÇÃO

ARTIFICIAL INTELLIGENCE AGENTS FOR REFLEXIVE THEMATIC ANALYSIS WITHOUT PROGRAMMING

Rafael Cardoso Sampaio[1] ORCID: 0000-0001-5176-173X [1] Universidade Federal de Pernambuco, Departamento de Ciência Política. Recife, Pernambuco, Brasil.

Cómo citar: Sampaio RC. Artificial intelligence agents for reflexive thematic analysis without programming. Online Braz J Nurs. 2026;25(Suppl1):e20267166. https://doi.org/10.17665/1676-4285.20267166

RESUMO

Objetivo: descrever uma *skill* para análise temática reflexiva e mostrar que pesquisadores sem formação em programação podem usar agentes de inteligência artificial na análise de dados qualitativos de modo simples, confiável, transparente e replicável. **Método:** estudo metodológico. A *skill* foi construída a partir das seis fases da análise temática reflexiva e dos riscos documentados do uso de inteligência artificial generativa. Foi aplicada a um corpus real de doze sessões de grupos focais sobre sentidos de democracia e proteção social e avaliada em teste controlado com 126 execuções cegas. **Resultados:** a *skill* organiza a análise em sete fases, da decisão metodológica ao relatório final, com verificação literal das citações e bloqueio de atalhos. No corpus real, produziu o fluxo completo com cobertura total. No teste cego, a condição com a *skill* teve maior aderência aos critérios metodológicos. **Conclusão:** agentes de inteligência artificial podem ser usados sem programação, com método e rastreabilidade, desde que a interpretação e a responsabilidade sejam humanas.
Descritores: Inteligência Artificial Generativa; Pesquisa Qualitativa; Análise de Dados; Reprodutibilidade dos Testes.

ABSTRACT

Objective: To describe a *skill* for reflexive thematic analysis and to show that researchers without programming training can use artificial intelligence agents to analyze qualitative data in a simple, reliable, transparent, and replicable way. **Method:** Methodological study. The *skill* was developed based on the six phases of reflexive thematic analysis and on the documented risks of using generative artificial intelligence. It was applied to a real corpus of twelve focus group sessions on the meanings of democracy and social protection and evaluated in a controlled test with 126 blind runs. **Results:** The *skill* organizes the analysis into seven phases, from the methodological decision to the final report, with literal verification of quotations and checks that block shortcuts. In the real corpus, it carried out the complete workflow with full coverage. In the blind test, the condition with the *skill* showed greater adherence to the methodological criteria. **Conclusion:** Artificial intelligence agents can be used without programming, in a methodical and traceable way, provided that interpretation and responsibility remain human.
Descriptors: Generative Artificial Intelligence; Qualitative Research; Data Analysis; Reproducibility of Results.

RESUMEN

Objetivo: describir una *skill* para el análisis temático reflexivo y mostrar que investigadores sin formación en programación pueden utilizar agentes de inteligencia artificial en el análisis de datos cualitativos de modo sencillo, fiable, transparente y replicable. **Método:** estudio metodológico. La *skill* se construyó a partir de las seis fases del análisis temático reflexivo y de los riesgos documentados del uso de la inteligencia artificial generativa. Se aplicó a un corpus real de doce sesiones de grupos focales sobre los sentidos de la democracia y la protección social y se evaluó en una prueba controlada con 126 ejecuciones ciegas. **Resultados:** la *skill* organiza el análisis en siete fases, desde la decisión metodológica hasta el informe final, con verificación literal de las citas y bloqueo de atajos. En el corpus real, produjo el flujo completo con cobertura total. En la prueba ciega, la condición con la *skill* mostró mayor adherencia a los criterios metodológicos. **Conclusión:** los agentes de inteligencia artificial pueden utilizarse sin programación, con método y trazabilidad, siempre que la interpretación y la responsabilidad sean humanas.
Descriptores: Inteligencia Artificial Generativa; Investigación Cualitativa; Análisis de Datos; Reproducibilidad de los Resultados.

Editores:
Rosimere Ferreira Santana (ORCID: 0000-0002-4593-3715)
Audrey Vidal Pereira (ORCID: 0000-0002-6570-9016)

Editora:
Escola de Enfermagem Aurora de Afonso Costa – UFF
Rua Dr. Celestino, 74 – Centro, CEP: 24020-091 – Niterói, RJ, Brasil
Correo electrónico de la revista: objn.cme@id.uff.br

Autor correspondiente:
Rafael Cardoso Sampaio
Correo electrónico: cardososampaio@gmail.com

Lo que ya se sabe:

- La IA generativa ya se usa en el análisis cualitativo sin documentación adecuada; los *chatbots* pueden fabricar citas y generar temas sin trazabilidad.
- Los flujos basados en programación ofrecen más control y reproducibilidad, pero exigen conocimientos técnicos que la mayoría de los investigadores cualitativos no tiene.
- El análisis temático reflexivo exige decisiones interpretativas humanas, y faltaba una forma sencilla y auditable de conducirlo con agentes de IA.

Lo que este artículo aporta:

- El artículo desarrolla una *skill* abierta para que investigadores sin programación conduzcan el análisis temático reflexivo con agentes de IA, en lenguaje natural.
- La *skill* verifica que cada cita exista realmente en las transcripciones originales, lo que impide que la IA invente citas. Además, registra cada decisión, haciendo el análisis auditable y contestable por terceros.
- En la prueba ciega de 126 ejecuciones, la *skill* mostró mayor adherencia a los criterios metodológicos. Aun así, esto no exime de la validación humana.

INTRODUCCIÓN

Un investigador termina la recolección de veinte entrevistas y necesita analizar el material; entonces abre una plataforma de inteligencia artificial generativa (IAG), pega fragmentos de las transcripciones y recibe una lista de temas que parecen terminados. La respuesta es fluida y convincente, pero también es difícil de verificar, porque no hay registro de cómo se construyeron los temas, ni indicación de dónde se encuentra cada idea en el material, ni garantía de que las intervenciones citadas existan[1]. La inquietud ante ese resultado está en un hallazgo que no puede rastrearse y en un dato que no puede defenderse[2].

El uso de la inteligencia artificial generativa en la investigación cualitativa creció más rápido que la documentación de ese uso[3]. Los protocolos para conducir el análisis temático con el *chatbot* ChatGPT ya circulan en revistas revisadas por pares[4], y las comparaciones controladas entre modelos e investigadores humanos muestran resultados cercanos en la codificación, con diferencias de peso en la interpretación[5], y la mayoría de los sistemas registra un predominio de arreglos híbridos, en los que la última palabra es humana[6]. Parte de los riesgos está en la fabricación o el pulido de citas[2]; en la tendencia del modelo a coincidir con la hipótesis del investigador, conocida como sicofancia (*sycophancy*) o, simplemente, adulación[6]; en textos fluidos en la superficie que disimulan la ausencia de interpretación[7]; y en el borrado silencioso de vivencias poco frecuentes[6]. En el mismo período, el área pasó a producir ejemplos de colaboración estructurada entre humanos y máquinas, con protocolos explícitos y registro de las decisiones[8].

El *chatbot* es fácil de manejar, pero frágil en el control metodológico, porque cada conversación empieza de cero. Los flujos de trabajo basados en programación suelen ofrecer mayor control y reproducibilidad, pero exigen conocimientos técnicos y alejan justamente a quienes más necesitarían documentación asistida[9]. Existe, sin embargo, un camino intermedio, aún poco explorado en la literatura y desconocido para la mayoría de los investigadores cualitativos.

Este camino se apoya en el uso de agentes de inteligencia artificial (IA), que se sustentan en una estructura que involucra un modelo de IAG y le permite usar herramientas, seguir reglas y ejecutar tareas de forma más controlada. El agente puede acceder a archivos y carpetas, buscar en la web, ejecutar código, usar API, consultar bases de datos, crear y editar documentos, probar resultados, comparar fuentes, registrar las etapas del trabajo y corregir errores antes de entregar una respuesta. Esta estructura se conoce como *harness*[9]. En la práctica, esos arneses pueden usarse en programas de fácil instalación, que funcionan con suscripciones a los modelos de IAG, como Antigravity, de Google; Claude Code, de Anthropic; y Codex, de OpenAI.

A diferencia de los *chatbots*, los agentes de IA pueden ejecutar tareas largas y complejas porque planifican el trabajo, lo dividen en etapas y gestionan el límite de información que el modelo de IAG logra mantener presente a la vez. Cuando ese límite se acerca, el agente puede resumir o registrar lo que ya se hizo, liberar espacio y continuar la tarea sin necesidad de empezar de nuevo.

Una de las maneras de realizar tareas largas es, precisamente, el uso de *skills* (habilidades) diseñadas para una tarea determinada. Una *skill* es un procedimiento reutilizable empaquetado, por lo general como un archivo o una carpeta de instrucciones en texto simple, que el agente carga cuando reconoce cierto tipo de tarea. Puede contener desde orientaciones sencillas hasta un flujo de trabajo complejo, con varias etapas, criterios de decisión, herramientas que deben usarse, formas de verificar resultados, reglas para manejar errores e instrucciones sobre cómo registrar cada etapa. En lugar de limitarse a decirle al agente qué hacer, una *skill* puede especificar cómo ejecutar todo el proceso. Como el protocolo es abierto, cualquier investigador puede publicar, adaptar y someter a la crítica de sus pares el procedimiento que utilizó.

El análisis temático reflexivo, formulado por Braun y Clarke en 2006[10]y reformulado en 2019[11], ofrece un buen caso para demostrar esa posibilidad. El método se organiza en seis fases y sostiene que los temas son construidos por el analista a partir de patrones de significado, y no descubiertos ya listos en el material. La recepción brasileña muestra que esa distinción no siempre llega a la práctica[12]. El análisis temático reflexivo tiene criterios claros y fases definidas, y exige del investigador decisiones que no pueden delegarse, lo que lo convierte en un buen caso para mostrar lo que hace una *skill*.

La *skill* analise-tematica-reflexiva es un paquete abierto que conduce el análisis temático reflexivo en siete fases, desde la decisión metodológica hasta el informe final. El objetivo de este artículo es describir la construcción de esa *skill* y la demostración de su uso en un corpus real de investigación cualitativa, mostrando cómo un investigador sin formación en programación puede analizar datos cualitativos con agentes de inteligencia artificial. La ganancia científica está en lograr que el análisis sea sencillo, fiable,

transparente y replicable, manteniendo la interpretación como tarea humana.

MÉTODO

Este estudio describe el desarrollo de una *skill* y la demostración de su uso en un corpus real de investigación cualitativa. La *skill* nació del cruce entre las seis fases del análisis temático reflexivo y los riesgos documentados del uso de la inteligencia artificial generativa[2,6,7]. Cada fase del método se convirtió en una etapa con salida obligatoria y una verificación que bloquea atajos, como saltarse la lectura inicial del material, aceptar códigos vagos, citar un extracto no verificado o tratar un resumen de asunto como tema.

La *skill* organiza el análisis en siete fases, descritas en la Tabla 1. En cada fase, el agente propone y el investigador decide, y la separación entre propuesta y decisión queda registrada en un documento propio. Las fases F1 a F6 corresponden a las seis fases de Braun y Clarke[10-11]; la fase F0 es un añadido de gobernanza, que obliga al investigador a registrar las decisiones antes de codificar, y la separación entre contestación y definición, en las fases F4 y F5, explicita en etapas distintas lo que el método trata en dos momentos.

Tabla 1 – Fases de la *skill*, lo que el investigador decide y lo que el agente produce

Fase	El investigador decide	El agente produce
F0. Decisiones metodológicas	Pregunta, corpus, posición teórica y cuánto delegar	Plan del análisis, con límites declarados
F1. Lectura inicial	Qué merece registro	Anotaciones por fuente e inventario de lo leído
F2. Códigos	Qué códigos entran y cuáles salen	Códigos anclados en extractos localizados
F3. Temas candidatos	Qué patrón de significado organiza cada tema	Mapa de temas por significado, con residuos aparte
F4. Contestación	Qué lectura alternativa se acepta o se rechaza	Lectura rival y examen de contradicciones
F5. Definición y denominación	Nombre y límite de cada tema	Pasaporte de cada tema, con genealogía y extractos
F6. Informe	Qué puede afirmarse	Informe con declaración de uso de IA y estatus del análisis

Fuente: elaborado por el autor, 2026.

La *skill* deja que el investigador elija cuánto avanza la máquina por sí sola. En la vía más libre, el agente recorre las siete fases y declara la ausencia de validación humana. En otra, se detiene en puntos definidos para su verificación. Existe además la posibilidad de que el investigador entregue anotaciones, códigos o temas iniciales, que el agente contrasta con el corpus. Y existe la vía paralela, en la que el agente lee todo antes de ver la lectura humana, y ambas lecturas se comparan al final, sin que la concordancia se trate como validación. El nombre análisis temático reflexivo se reserva para el recorrido con analista humano activo; la vía más libre produce material de exploración auditable, y el estatus final registra esa condición con claridad.

El uso de la *skill* no exige escribir código, pero sí instalar un programa de agente y copiar una carpeta de archivos, operaciones básicas que cualquier usuario de computadora realiza. Para quien empieza, los programas con interfaz visual, como Antigravity 2.0, Codex/ChatGPT Work o Claude Code/Cowork, son el camino más corto[9]. La *skill* es abierta y está disponible en un repositorio público[6]. Al investigador le basta pedirle al agente que instale la *skill*, facilitándole la URL, y este la copiará en la carpeta de habilidades del programa. A continuación, debe abrir el agente en una carpeta de su propia computadora, donde estará el material que se va a analizar, como transcripciones de entrevistas y grupos focales. Después, basta hacer la solicitud en lenguaje natural, como se le pediría a un asistente de investigación. Siguiendo la *skill*, el agente preguntará qué tipo de análisis se va a realizar y seguirá las etapas solicitadas dentro de ella.

Para demostrar el uso, la *skill* se aplicó a un corpus real de doce sesiones de grupos focales realizados en distintas regiones de Brasil, con 154.573 palabras, sobre los sentidos de la democracia y la protección social entre participantes de programas de políticas públicas, en la aplicación Codex de OpenAI, con el modelo ChatGPT Luna 6 y nivel de esfuerzo extra alto. La solicitud se hizo en lenguaje natural, con la pregunta de investigación y la ruta autónoma, en la que el agente recorre las fases sin paradas de verificación; no hubo participación humana durante el flujo, y por eso el estatus final registra la ausencia de validación humana. El corpus utilizado se basó en grupos focales realizados a lo largo de 2024 en 11 capitales del país, divididos entre cuatro franjas de ingresos, definidas en salarios mínimos: 0-2, 2-5, 5-10 y 10-20. Más específicamente, la *skill* trabajó con un corpus derivado (basado solo en Belém, Recife y Curitiba) y anonimizado respecto del original, de modo que se impidiera reconstruir el contexto original e identificar a los participantes.

La *skill* también se sometió a una evaluación comparativa controlada, descrita en la parte final de los resultados, que contrastó el mismo modelo con y sin la *skill* en 126 ejecuciones ciegas, organizadas como veintiuna situaciones de falla metodológica repetidas tres veces en cada condición. Las situaciones de falla metodológica son “escenarios trampa”, casos diseñados para probar si el protocolo rechaza los atajos. Por ejemplo: saltarse la lectura inicial y codificar directamente, tratar un resumen de asunto como

tema, publicar una cita que no existe en el corpus, exponer identificadores de participantes, inventar el consentimiento para el uso de IA, armonizar intervenciones para forzar una conclusión, etc. Cada escenario especifica el comportamiento erróneo esperado sin la *skill* y el comportamiento exigido con la *skill*. La evaluación y sus límites forman parte del repositorio de la *skill*[6], y se verificó la portabilidad entre dos agentes de IA distintos.

RESULTADOS

Lo que recibe el investigador

La Tabla 2 muestra, en lenguaje natural, la solicitud que dio inicio a la demostración.

Tabla 2 – Ejemplo de solicitud al agente y lo que este entrega

El investigador pide	El agente entrega
Analiza las transcripciones de esta carpeta sobre los sentidos de la democracia y la protección social, mediante el análisis temático reflexivo, en la ruta autónoma.	Siete documentos, uno por fase, con informe final, apéndice de auditoría y estatus del análisis

Fuente: elaborado por el autor, 2026.

De la solicitud salieron siete documentos, uno por fase, además del informe final y del apéndice de auditoría. El recorrido completo llevó menos de una hora en esta demostración, tiempo estimado de manera informal, frente a semanas de trabajo manual. El uso de la IAG en el método cualitativo, sin embargo, no puede atarse exclusivamente a la ganancia de eficiencia: es necesario verificar también la calidad de la propia evaluación.

Antes del detalle, una nota sobre la correspondencia de fases. Cada fase de la *skill* corresponde a una de las seis fases del análisis temático reflexivo de Braun y Clarke: la Fase 1 es la familiarización con los datos; la Fase 2, la generación de códigos iniciales; las Fases 3 y 4, la búsqueda y la revisión de temas; la Fase 5, la definición y la denominación; y la Fase 6, la elaboración del informe. La Fase 0 es la excepción: no existe en Braun y Clarke, sino que es un añadido de gobernanza para registrar las decisiones antes de que el flujo comience.

El primer documento registra las decisiones del análisis, como pregunta de investigación, corpus, unidad de análisis, posición teórica, orientación inductiva/deductiva y cuánto se delegaría en cada fase. En la demostración, la pregunta indagaba cómo participantes de distintas capitales y franjas de ingresos construyen sentidos sobre la democracia, el Estado, la corrupción, la protección social y la vida moral, y cómo la posición socioeconómica atraviesa esos sentidos; la posición teórica se declaró contextualista crítico-interpretativa, y la ciudad y los ingresos entraron como atributos de contraste, no como temas. La *skill* no avanza mientras ese registro no exista.

La Fase 1, lectura inicial, corresponde a la primera fase de Braun y Clarke, de familiarización con los datos, y viene antes de cualquier código. El agente inventarió el corpus (13 archivos: las 12 transcripciones más el índice de sesiones), recorrió las doce sesiones repartidas entre Belém, Recife y Curitiba en las cuatro franjas de ingresos y produjo anotaciones por sesión, con el perfil del grupo, sorpresas y pasajes de interés, sin generar un solo código.

La Fase 2, códigos, corresponde a la segunda fase de Braun y Clarke, de generación de códigos iniciales, y ancló cada código en extractos con localización precisa. Fueron 54 códigos, de C01 a C54, cada uno con significado, extracto representativo, localizador en el corpus y nota interpretativa. A modo de ejemplo, el C01, "Democracia como procedimiento electoral y libertad abstracta", se sustentó en "democracia es tener libertad de expresión, votar por quien uno quiere, libre albedrío" (BEL-0-2, línea 16); el C04, "Autoritarismo como imposición y 'no tener voz'", en "no te da voz, donde no tienes derechos, opciones" (BEL-5-10, línea 13); y los C42 y C43, "Programas sociales como necesarios con control" y "Programas sociales como 'acomodación' y manipulación", disputaron el mismo material de la protección social. La verificación de la distribución confirmó que los códigos cubrían las doce sesiones.

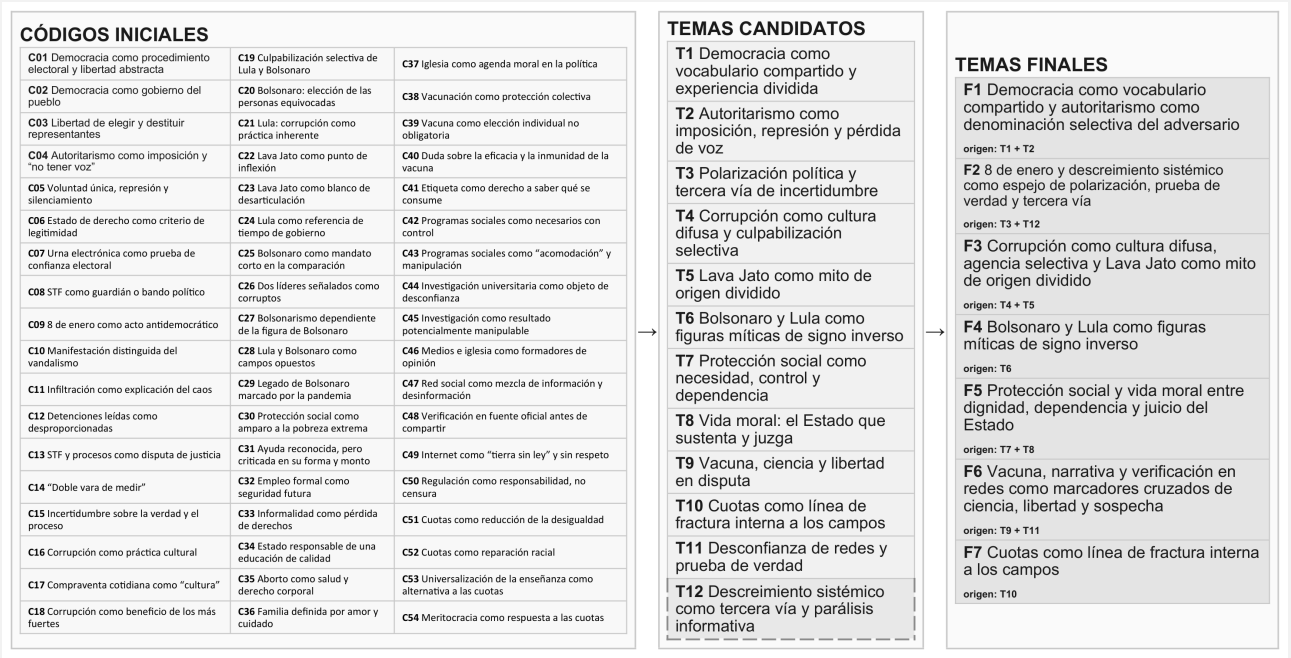
La Fase 3, temas candidatos, corresponde a las fases de búsqueda inicial de temas de Braun y Clarke, y construyó esos elementos a partir de patrones de significado compartido, y no por agrupación por asunto. Cada tema candidato debió cargar un concepto organizador y una afirmación interpretativa, nunca una etiqueta de dominio; los ingresos y la ciudad se rechazaron como tema, porque entran como atravesamiento, no como clave. Del mapa de 54 códigos salieron 11 candidatos, de T1 a T11. El primero, "Democracia como vocabulario compartido y experiencia dividida", argumentaba que casi todos definen la democracia con la misma fórmula (voto, libertad, elección), pero que el sentido práctico se bifurca según la experiencia de la derrota electoral, de la confianza institucional y de los ingresos. Lo que no encajaba en ningún candidato fue a una carpeta provisional de sobrantes, en lugar de ser descartado.

La Fase 4, de contestación, no aceptó el mapa temático sin examen. El agente integrador exigió una lectura alternativa plausible y una prueba de contradicciones y sobrantes, y la lectura rival llegó en forma de hipótesis: el organizador maestro de los sentidos no son los ingresos ni la ciudad, sino la posición electoral declarada en las sesiones pareadas, de modo que los ingresos y la ciudad seguirían siendo el contexto de producción de la experiencia material. La rival no derrotó a ningún tema, y esa no era su función; su función era forzar el examen. Forzó algunos cambios: el T3 se revisó para incorporar la tercera vía, el "no sé si fue verdad"; el T9 y el T10 se fortalecieron con casos de disidencia; el T11 se dividió, porque mezclaba dos movimientos distintos, el de construir narrativa y el de desconfiar de la verificación; y nació un tema nuevo, el T12, "Descreimiento sistémico como tercera vía y parálisis informativa".

Tener 12 temas muestra amplitud analítica, pero dificultad de síntesis. Por eso, en la propia *skill*, establecí como regla de oro apuntar a entre 4 y 8 temas, si el corpus lo permite. La agregación no fue aritmética, pues cada fusión debió demostrar que dos temas hablaban del mismo fenómeno. Así, el T1 y el T2 se fusionaron, porque nombrar la democracia y nombrar el autoritarismo son el mismo acto de posicionamiento en el léxico político; el T3 y el T12 se fusionaron, porque el descreimiento sistémico es la

cara de la misma polarización que expuso el 8 de enero, y la tercera vía del T3 era la parálisis del T12 en acto; el T4 y el T5 se fusionaron, porque el Lava Jato es el mito de origen de la misma cultura difusa de corrupción que culpabiliza de forma selectiva; el T7 y el T8 se fusionaron, porque la protección social y la vida moral son las dos caras de la misma relación, la del Estado que sustenta y juzga a la misma población; el T9 y el T11 se fusionaron, porque vacuna, narrativa y verificación son variaciones de la misma sospecha epistémica, y la desconfianza hacia las redes migró al tema de la prueba de verdad. El T6 y el T10 resistieron la fusión, por saturación y distinción demostrada. Los doce temas iniciales se convirtieron en siete (Figura 1), una aproximación más sustantiva y más fácil para contar de nuevo la historia contenida en los temas, tal como proponen Braun y Clarke.

Figura 1 – Flujo de la construcción de los temas, de los códigos iniciales a los temas candidatos y a los temas finales



Fuente: elaborado por el autor, 2026.

La Fase 5, de definición y denominación, según el método de las autoras, continuó con los siete temas restantes. Cada tema debió presentar ahora nombre, concepto organizador, afirmación interpretativa, genealogía (de los 12 temas precedentes hasta los códigos), extractos clave para la interpretación, la lectura rival rechazada y la zona de incertidumbre. Antes de eso, los extractos pasaron por una verificación literal contra el texto original, y el resultado confirmó que las 40 citas candidatas realmente existen en el corpus original, pero ninguna de ellas se publicó en su forma literal, pues fueron parafraseadas para impedir la identificación de los participantes.

El proceso muestra que los siete temas no fueron descubiertos en el corpus; fueron construidos, y construidos en medio de la disputa. Los temas finales fueron estos: Democracia como vocabulario compartido y autoritarismo como denominación selectiva del adversario; 8 de enero y descreimiento sistémico como espejo de polarización, prueba de verdad y tercera vía; Corrupción como cultura difusa, agencia selectiva y Lava Jato como mito de origen dividido; Bolsonaro y Lula como figuras míticas de signo inverso; Protección social y vida moral entre dignidad, dependencia y juicio del Estado; Vacuna, narrativa y verificación en redes como marcadores cruzados de ciencia, libertad y sospecha; y Cuotas como línea de fractura interna a los campos.

La Fase 6, de informe, que corresponde a la sexta fase de Braun y Clarke, fue la etapa de amarre. El texto contó la historia analítica en un cuerpo principal y un apéndice auditable. En el cuerpo principal, la pregunta de investigación amarra el corpus a la narrativa de los siete temas, cada uno con extractos ancla interpretados, variaciones entre fuentes y la rival rechazada registrada junto a la tesis que venció. En el apéndice se reunieron las fichas completas y contestables de cada tema final, con su genealogía, los extractos ancla verificados, la rival interpretativa, las evidencias decisivas, los residuos, los límites y las incertidumbres del análisis. También se incluyeron el manifiesto de uso de IA, la matriz de delegación por fase, los registros de interacción, el mapa humano, la auditoría de cobertura, la verificación de los extractos, los puntos de control y las decisiones adversariales.

Un mecanismo adicional para garantizar la calidad de la codificación está en la verificación que busca cada fragmento citado en el texto original y separa lo que es idéntico de lo que exige normalización, y solo lo idéntico se pone entre comillas. Sin esa comprobación, parte intrínseca de la *skill*, la IA tendría más libertad para inventar y llevar fragmentos incorrectos o imprecisos a la evaluación final. Esta regla vale para el análisis hecho sobre el corpus original; en esta publicación, los extractos fueron parafraseados para impedir la identificación de los participantes.

Evaluación del método

La *skill* se sometió a una prueba comparativa ciega con 126 ejecuciones, sesenta y tres sin la *skill* y sesenta y tres con ella. Las veintiuna situaciones de la prueba simulan los atajos que el protocolo debe rechazar, como saltarse la lectura inicial o tratar un resumen de asunto como tema. Un agente de IA basado en el modelo GLM 5.2 de Zhipu AI, que no participó en la generación de los materiales, aplicó una evaluación de 14 criterios a cada resultado (*output*), verificando aspectos como la presencia de extractos literales y la separación entre propuesta y decisión. En la evaluación ciega, la condición con la *skill* obtuvo mayor adherencia a los criterios metodológicos y fue evaluada por el agente de IA como la que obtuvo mejores resultados. El resultado es favorable, sin establecer validez, porque ninguna prueba sustituye a la evaluación humana.

DISCUSIÓN

La *skill* reduce la distancia entre el chat simple y el flujo de trabajo programático ("agéntico") sin exigir programación al investigador. El protocolo del análisis temático reflexivo, distribuido en siete fases con verificaciones, rutas y contestación, se vuelve comportamiento observable del agente, y cada fase produce insumos que pueden examinarse, auditarse y contestarse. La ganancia está en hacer que el proceso sea reconstruible por terceros, reforzando principios de replicabilidad, sin sustituir el juicio humano.

En mi opinión, el uso de la inteligencia artificial en la investigación cualitativa ya ocurre en investigaciones, laboratorios y orientaciones, sin enunciarse y sin documentación adecuada. La *skill* busca responder a ese problema al organizar criterios, fases y verificaciones en un artefacto único, que cualquier investigador puede inspeccionar, adaptar y criticar. Un evaluador que discrepa de un análisis puede instalar la *skill*, aplicarla con parámetros distintos e identificar dónde diverge la interpretación; un tutor puede usarla como objeto de enseñanza. En la práctica, el usuario puede incluso modificar la *skill* como desee, o hasta diseñar una desde cero simplemente pidiéndoselo al agente de IA. En esta prueba inicial, opté por hacer una evaluación solo con dos agentes de IA, pero crear criterios para evaluar la calidad por validadores humanos es, sin duda, un camino a seguir.

La promesa de transparencia encuentra terreno fértil en la investigación cualitativa en salud y enfermería, campo en el que las exigencias de reporte transparente crecen[3,13] y el análisis temático de datos como entrevistas y grupos focales se viene consolidando en el área[14].

Es necesario admitir, desde el principio, ciertos límites para tales análisis. La inteligencia artificial no tiene vivencia, juicio situado ni reflexividad humana, y la verificación de citas no garantiza la calidad interpretativa; o más aún, el inventario de archivos generados por la *skill* confirma la identidad del documento, no lo que el agente leyó en él, y la concordancia entre agentes no equivale a la triangulación humana. Los costos y las suscripciones siguen creando barreras, y los planes caros pesan más para estudiantes de posgrado y docentes brasileños[9]. Los datos sensibles exigen un cuidado redoblado, ya que las intervenciones de participantes procesadas por servicios de terceros imponen la verificación del consentimiento original y de las normas de protección de datos personales, y el procesamiento en servicios extranjeros implica transferencia internacional de datos, que el investigador debe evaluar según los términos de cada proveedor. La decisión sobre lo que puede salir de la computadora pertenece al investigador y al comité de ética[15].

La ruta autónoma, en la que el agente recorre las fases sin intervención, sigue siendo experimental y no declara equivalencia con la codificación. Hace visible lo que ocurre cuando la máquina ejecuta el protocolo completo y produce material que otro investigador puede auditar, contestar o usar como lectura inicial o incluso como lectura rival.

CONCLUSIÓN

Este texto buscó presentar una forma bastante sencilla de ejecutar un análisis temático automatizado, bastando instalar el programa de agente, descargar la *skill*, abrir la carpeta con el material que se va a analizar y pedir el análisis en lenguaje natural. Este camino me parece innovador y prometedor, puesto que facilita el acceso de cualquier investigador sin exigir conocimientos de programación, al mismo tiempo que genera materiales para auditar, cuestionar o incluso fundamentar el análisis. Quien antes pegaba fragmentos en el *chatbot* ahora puede pedir el análisis completo en lenguaje natural y mostrar, documento por documento, de dónde salió cada tema y cada intervención.

La demostración en un corpus real mostró que el flujo completo es posible, con cobertura total y cada cita contrastada con el original. La prueba ciega es compatible con la reducción de desvíos metodológicos en relación con el mismo modelo sin orientación, y la portabilidad mostró que el paquete no depende de un único proveedor. Ninguna de esas evidencias autoriza a afirmar equivalencia humana ni validez automática; sustentan la viabilidad, la transparencia y la auditabilidad, propiedades que la investigación cualitativa necesita defender en la era de la inteligencia artificial.

El límite obvio en esta etapa es la falta de validación con uno o más investigadores humanos y con análisis temáticos de calidad, algo que se realizará en las próximas etapas de la investigación. La validación por dos agentes distintos de IA también tiene límites, pues puede evidenciar apenas que esas tecnologías presentan sesgos similares; sin embargo, diversos experimentos personales y el experimento de validación indican un mejor desempeño en la codificación de aspectos cualitativos.

El valor científico de un agente cualitativo está en cuánto permite que otros reconstruyan, contesten y transformen lo que este produjo. La interpretación final y la responsabilidad por cada afirmación siguen siendo humanas. La herramienta debe ayudar al investigador a pensar mejor, nunca a pensar en su lugar.

CONFLICTO DE INTERESES

El autor declara no tener conflictos de intereses.

FINANCIACIÓN

El presente trabajo se realizó con el apoyo de una beca de productividad del CNPq.

USO DE INTELIGENCIA ARTIFICIAL

La *skill* descrita en este artículo fue desarrollada y evaluada con el uso de herramientas de inteligencia artificial generativa. ChatGPT Luna 6 se utilizó para las pruebas de la *skill* de análisis temático y también para correcciones y mejoras en el texto, además de haber auxiliado en la traducción. La redacción de este manuscrito contó con la asistencia de esas herramientas, con revisión y edición final por el autor, quien asume íntegramente la responsabilidad por el contenido.

REFERENCIAS

1. Paulus TM, Marone V. "In minutes instead of weeks": discursive constructions of generative AI and qualitative data analysis. Qual Inq. 2025;31(5):395-402. https://doi.org/10.1177/10778004241250065
2. Jowsey T, Stapleton P, Campbell S, Davidson A, McGillivray C, Maugeri I, et al. Frankenstein, thematic analysis and generative artificial intelligence: quality appraisal methods and considerations for qualitative research. PLoS One. 2025;20(9):e0330217. https://doi.org/10.1371/journal.pone.0330217. PMID: 40911617.
3. Kempny C, Frings J, Rust P, Meister S, Fehring L. The use and methodological reporting of large language models in qualitative research: a scoping review. BMC Med Res Methodol. 2026;26(1):137. https://doi.org/10.1186/s12874-026-02913-1. PMID: 42298426.
4. Naeem M, Smith T, Thomas L. Thematic analysis and artificial intelligence: a step-by-step process for using ChatGPT in thematic analysis. Int J Qual Methods. 2025;24. https://doi.org/10.1177/16094069251333886
5. Wachinger J, Bärnighausen K, Schäfer LN, Scott K, McMahon SA. Prompts, pearls, imperfections: comparing ChatGPT and a human researcher in qualitative data analysis. Qual Health Res. 2025;35(9):951-66. https://doi.org/10.1177/10497323241244669. PMID: 38775392.
6. Sampaio R. Análise temática reflexiva: skill para agentes de codificação [software]. 2026 [citado 2026 Ago 24]. Disponible en: https://github.com/labiia-lab/analise-tematica-reflexiva
7. Dellafiore F, Saba A, Collaro C, Artioli G. Artificial intelligence in qualitative research: insights from experts via reflexive thematic analysis. Qual Health Res. 2026;36(2-3):145-65. https://doi.org/10.1177/10497323251389800. PMID: 41351219.
8. Jaeger AA, Costa AP. Análise de dados guiada pelo modelo abductivai: um estudo de colaboração humano-IA. Rev Pesqui Qual. 2026;14(42):168-97. https://doi.org/10.33361/RPQ.2026.v.14.n.42.1629
9. Sampaio RC. Agentes de IA na pesquisa acadêmica: um tutorial para pesquisadores sem programação. Salvador: Edufba; 2026. En prensa.
10. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol. 2006;3(2):77-101. https://doi.org/10.1191/1478088706qp063oa
11. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health. 2019;11(4):589-97. https://doi.org/10.1080/2159676X.2019.1628806
12. Souza LK. Pesquisa com análise qualitativa de dados: conhecendo a análise temática. Arq Bras Psicol [Internet]. 2019 [citado 2026 Ago 24];71(2):51-67. Disponible en: http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1809-52672019000200005
13. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. Int J Qual Methods. 2017;16(1):1-13. https://doi.org/10.1177/1609406917733847
14. Dias EG, Mishima SM. Análise temática de dados qualitativos: uma proposta prática para efetivação. Sustinere. 2023;11(1):402-11. https://doi.org/10.12957/sustinere.2023.71828
15. Sampaio RC, Sabbatini M, Limongi R. Diretrizes para o uso ético e responsável da inteligência artificial generativa: um guia prático para pesquisadores [Internet]. São Paulo: Intercom; 2024 [citado 2026 Ago 24]. Disponible en: https://www.portcom.intercom.org.br/ebooks/arquivos/livro-diretrizes-ia.pdf

CONTRIBUCIÓN DE AUTORÍA

Diseño del estudio: Sampaio RC.
Obtención de datos: Sampaio RC.
Análisis de datos: Sampaio RC.
Interpretación de datos: Sampaio RC.
El autor se responsabiliza de la redacción y la revisión crítica del contenido intelectual, de la aprobación de la versión final publicada y de todos los aspectos éticos, legales y científicos relacionados con la exactitud y la integridad del estudio.

© 2026 Autores. Los autores conservan los derechos de autor y conceden a la Online Brazilian Journal of Nursing el derecho de primera publicación. Este artículo está licenciado bajo la licencia Creative Commons Atribución 4.0 Internacional (CC BY 4.0).