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

Submissão: 03-Set-2026
Aprovado: 20-Out-2026

AGENTES DE INTELIGÊNCIA ARTIFICIAL PARA ANÁLISE TEMÁTICA REFLEXIVA SEM PROGRAMAÇÃO

ARTIFICIAL INTELLIGENCE AGENTS FOR REFLEXIVE THEMATIC ANALYSIS WITHOUT PROGRAMMING

AGENTES DE INTELIGENCIA ARTIFICIAL PARA EL ANÁLISIS TEMÁTICO REFLEXIVO SIN PROGRAMACIÓN

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

Como 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
E-mail da revista: objn.cme@id.uff.br

Autor correspondente:
Rafael Cardoso Sampaio
E-mail: cardososampaio@gmail.com

O que já se sabe:

- A IA generativa já é usada na análise qualitativa sem documentação adequada; *chatbots* podem fabricar citações e gerar temas sem rastreabilidade.
- Fluxos de trabalho com programação dão mais controle e reprodutibilidade, mas exigem conhecimento técnico que a maioria dos pesquisadores qualitativos não tem.
- A análise temática reflexiva exige decisões interpretativas humanas, e faltava uma forma simples e auditável de conduzi-la com agentes de IA.

O que este artigo acrescenta:

- O artigo desenvolve uma *skill* aberta que permite a pesquisadores sem formação em programação conduzir a análise temática reflexiva com agentes de IA, em linguagem natural.
- A *skill* verifica se cada fala citada existe de fato nas transcrições originais, o que impede a IA de inventar citações. Além disso, registra cada decisão, tornando a análise auditável e contestável por terceiros.
- No teste cego de 126 execuções, a *skill* teve maior aderência aos critérios metodológicos. Ainda assim, isso não dispensa a validação humana.

INTRODUÇÃO

Um pesquisador termina a coleta de vinte entrevistas e precisa analisar o material, então abre uma plataforma de inteligência artificial generativa (IAG), cola trechos das transcrições e recebe uma lista de temas que parecem prontos. A resposta é fluida e convincente, e também é difícil de conferir, porque não há registro de como os temas foram construídos, nem indicação de onde cada ideia se encontra no material, nem garantia de que as falas citadas existem[1]. O incômodo diante desse resultado está em um achado que não pode ser rastreado e em um dado que não pode ser defendido[2].

O uso de inteligência artificial generativa na pesquisa qualitativa cresceu mais rápido do que a documentação desse uso[3]. Protocolos para conduzir a análise temática com o *chatbot* ChatGPT já circulam em periódicos revisados por pares[4], e comparações controladas entre modelos e pesquisadores humanos mostram resultados próximos na codificação, com diferenças de monta na interpretação[5]e a maioria dos sistemas registra predomínio de arranjos híbridos, nos quais a palavra final é humana[6]. Parte dos riscos está na fabricação ou polimento de citações[2]; a tendência do modelo a concordar com a hipótese do pesquisador, conhecida como sicofância (*sycophancy)* ou simplesmente bajulação[6]; textos fluidos na superfície que disfarçam a ausência de interpretação[7]e apagamento silencioso de vivências raras[6]. No mesmo período, a área passou a produzir exemplos de colaboração estruturada entre humanos e máquinas, com protocolos explícitos e registro das decisões[8].

O *chatbot* é fácil de operar, mas é frágil no controle metodológico, porque cada conversa começa do zero. Fluxos de trabalho baseados em programação costumam ter maior controle e reprodutibilidade, mas exigem conhecimento técnico e afastam justamente quem mais precisaria de documentação assistida[9]. Existe, porém, um caminho intermediário, ainda pouco explorado na literatura e desconhecido pela maioria dos pesquisadores qualitativos.

Este caminho se baseia no uso de agentes de inteligência artificial (IA), que se apoiam em uma estrutura que envolve um modelo de IAG e permite que ele use ferramentas, siga regras e execute tarefas de forma mais controlada. Ele pode acessar arquivos e pastas, pesquisar na web, executar código, usar APIs, consultar bancos de dados, criar e editar documentos, testar resultados, comparar fontes, registrar etapas do trabalho e corrigir erros antes de entregar uma resposta. Essa estrutura é conhecida como *harness*[9]. Na prática, tais arreios podem ser usados em softwares de fácil instalação, que funcionam com assinaturas dos modelos de IAG, como o Antigravity, do Google; o Claude Code, da Anthropic; e o Codex, da OpenAI.

Diferentemente dos chatbots, agentes de IA podem executar tarefas longas e complexas porque planejam o trabalho, dividem-no em etapas e gerenciam o limite de informações que o modelo de IAG consegue manter em mente de uma só vez. Quando esse limite se aproxima, o agente pode resumir ou registrar o que já foi feito, liberar espaço e continuar a tarefa sem precisar recomeçar.

Uma das maneiras de realizar tarefas longas é justamente usando *skills* (habilidades) desenhadas para determinada tarefa. Uma *skill* é um procedimento reutilizável empacotado, em geral como um arquivo ou uma pasta de instruções em texto simples, que o agente carrega quando reconhece determinado tipo de tarefa. Ela pode conter desde orientações simples até um fluxo de trabalho complexo, com várias etapas, critérios de decisão, ferramentas a serem usadas, formas de verificar resultados, regras para lidar com erros e instruções sobre como registrar cada etapa. Em vez de apenas dizer ao agente o que fazer, uma *skill* pode especificar como executar todo o processo. Como o protocolo é aberto, qualquer pesquisador pode publicar, adaptar e submeter à crítica de pares o procedimento que usou.

A análise temática reflexiva, formulada por Braun e Clarke em 2006[10]e reformulada em 2019[11], oferece um bom caso para demonstrar essa possibilidade. O método organiza-se em seis fases e sustenta que os temas são construídos pelo analista a partir de padrões de significado, e não descobertos prontos no material. A recepção brasileira mostra que essa distinção nem sempre chega à prática[12]. A análise temática reflexiva tem critérios claros e fases definidas, e exige do pesquisador decisões que não podem ser delegadas, o que a torna um bom caso para mostrar o que uma *skill* faz.

A *skill* analise-tematica-reflexiva é um pacote aberto que conduz a análise temática reflexiva em sete fases, da decisão metodológica ao relatório final. O objetivo deste artigo é descrever a construção dessa *skill* e a demonstração de seu uso em corpus real de pesquisa qualitativa, mostrando como um pesquisador sem formação em programação pode analisar dados qualitativos com agentes de inteligência artificial. O ganho científico está em tornar a análise simples, confiável, transparente e replicável,

mantendo a interpretação como tarefa humana.

MÉTODO

Este estudo descreve o desenvolvimento de uma *skill* e a demonstração de seu uso em um corpus real de pesquisa qualitativa. A *skill* nasceu do cruzamento entre as seis fases da análise temática reflexiva e os riscos documentados do uso de inteligência artificial generativa[2,6,7]. Cada fase do método virou uma etapa com saída obrigatória e uma verificação que bloqueia atalhos, como pular a leitura inicial do material, aceitar códigos vagos, citar excerto não verificado ou tratar um resumo de assunto como tema.

A *skill* organiza a análise em sete fases, descritas na Tabela 1. Em cada fase, o agente propõe e o pesquisador decide, e a separação entre proposta e decisão fica registrada em documento próprio. As fases F1 a F6 correspondem às seis fases de Braun e Clarke[10-11]; a fase F0 é um acréscimo de governança, que obriga o pesquisador a registrar as decisões antes de codificar, e a separação entre contestação e definição, nas fases F4 e F5, explicita em etapas distintas o que o método trata em dois momentos.

Tabela 1 – Fases da *skill*, o que o pesquisador decide e o que o agente produz

Fase	O pesquisador decide	O agente produz
F0. Decisões metodológicas	Pergunta, corpus, posição teórica e quanto delegar	Plano da análise, com limites declarados
F1. Leitura inicial	O que merece registro	Anotações por fonte e inventário do que foi lido
F2. Códigos	Quais códigos entram e quais saem	Códigos ancorados em excertos localizados
F3. Temas candidatos	Qual padrão de significado organiza cada tema	Mapa de temas por significado, com resíduos à parte
F4. Contestação	Qual leitura alternativa é aceita ou recusada	Leitura rival e exame de contradições
F5. Definição e nomeação	Nome e limite de cada tema	Passaporte de cada tema, com genealogia e excertos
F6. Relatório	O que pode ser afirmado	Relatório com declaração de uso de IA e estatuto da análise

Fonte: elaborado pelo autor, 2026.

A *skill* deixa o pesquisador escolher o quanto a máquina anda sozinha. Na via mais livre, o agente percorre as sete fases e declara a ausência de validação humana. Em outra, ele para em pontos definidos para conferência. Há ainda a possibilidade de o pesquisador entregar anotações, códigos ou temas iniciais, que o agente confere contra o corpus. E existe a via paralela, na qual o agente lê tudo antes de ver a leitura humana, com as duas leituras comparadas ao final, sem que a concordância seja tratada como validação. O nome análise temática reflexiva é reservado ao percurso com analista humano ativo; a via mais livre produz material de exploração auditável, e o estatuto final registra essa condição com clareza.

O uso da *skill* não exige escrever código, mas exige instalar um programa de agente e copiar uma pasta de arquivos, operações básicas que qualquer usuário de computador realiza. Para quem começa, os programas com interface visual, como Antigravity 2.0, Codex/ChatGPT Work ou Claude Code/Cowork, são o caminho mais curto[9]. A *skill* é aberta e está disponível em repositório público[6]. Basta ao pesquisador pedir para o agente instalar a *skill* fornecendo-lhe a URL, e ele a copiará para a pasta de habilidades do programa. Em seguida, deve abrir o agente numa pasta do próprio computador do pesquisador, onde deverá estar o material a ser analisado, como transcrições de entrevistas e grupos focais. Depois, basta fazer o pedido em linguagem natural, como pediria a um assistente de pesquisa. Seguindo a *skill*, o agente perguntará o tipo de análise a ser feita e seguirá as etapas solicitadas dentro dela.

Para demonstrar o uso, a *skill* foi aplicada a um corpus real de doze sessões de grupos focais espalhados pelo Brasil, com 154.573 palavras, sobre os sentidos de democracia e proteção social entre participantes de políticas públicas no aplicativo Codex da OpenAI, usando o modelo ChatGPT Luna 6, com nível de esforço extra alto. O pedido foi feito em linguagem natural, com a pergunta de pesquisa e a rota autônoma, na qual o agente percorre as fases sem paradas de conferência; não houve participação humana durante o fluxo, e por isso o estatuto final registra validação humana ausente. O corpus utilizado foi baseado em grupos focais que aconteceram ao longo de 2024 em 11 capitais do país, divididos entre quatro faixas de renda, definidas em salários mínimos: 0-2, 2-5, 5-10 e 10-20. A *skill*, mais especificamente, trabalhou em um corpus derivado (baseado apenas em Belém, Recife e Curitiba) e anonimizado em relação ao original, de modo a impedir a reconstrução do contexto original e a identificação dos participantes.

A *skill* também foi submetida a uma avaliação comparativa controlada, descrita na parte final dos resultados, que confrontou o mesmo modelo com e sem a *skill* em 126 execuções cegas, organizadas como vinte e uma situações de falha metodológica repetidas três vezes em cada condição. Situações de falha metodológica são "cenários-armadilha", casos desenhados para testar se o protocolo recusa atalhos. Por exemplo: pular a leitura inicial e codificar direto, tratar um resumo de assunto como tema, publicar citação que não existe no corpus, expor identificadores de participantes, inventar consentimento para IA, harmonizar falas para forçar uma conclusão, etc. Cada cenário especifica o com-

portamento errado esperado *sem* a *skill* e o comportamento exigido *com* a *skill*. A avaliação e seus limites fazem parte do repositório da *skill*[6], e a portabilidade foi verificada entre dois agentes de IA distintos.

RESULTADOS

O que o pesquisador recebe

A Tabela 2 mostra, em linguagem natural, o pedido que deu início à demonstração.

Tabela 2 – Exemplo de pedido ao agente e o que ele entrega

O pesquisador pede	O agente entrega
Analise as transcrições desta pasta sobre os sentidos de democracia e proteção social, usando a análise temática reflexiva, na rota autônoma.	Sete documentos, um por fase, com relatório final, apêndice de auditoria e estatuto da análise

Fonte: elaborado pelo autor, 2026.

Do pedido saíram sete documentos, um por fase, além do relatório final e do apêndice de auditoria. O percurso completo levou menos de uma hora nesta demonstração, tempo estimado de forma informal, contra semanas de trabalho manual. O uso de IAG no método qualitativo, porém, não pode se atrelar exclusivamente ao ganho de eficiência: é preciso também verificar a qualidade da própria avaliação.

Antes do detalhamento, uma nota sobre a correspondência de fases. Cada fase da *skill* corresponde a uma das seis fases da análise temática reflexiva de Braun e Clarke: a Fase 1 é a familiarização com os dados; a Fase 2, a geração de códigos iniciais; as Fases 3 e 4, a busca e a revisão de temas; a Fase 5, a definição e a nomeação; e a Fase 6, a produção do relatório. A Fase 0 é a exceção: não existe em Braun e Clarke, é um acréscimo de governança para registrar as decisões antes de o fluxo começar.

O primeiro documento registra as decisões da análise, como pergunta de pesquisa, corpus, unidade de análise, posição teórica, orientação indutiva/dedutiva e quanto seria delegado em cada fase. Na demonstração, a pergunta indagava como participantes de diferentes capitais e faixas de renda constroem sentidos sobre democracia, Estado, corrupção, proteção social e vida moral, e como a posição socioeconômica atravessa esses sentidos; a posição teórica foi declarada como contextualista crítico-interpretativa, e cidade e renda entraram como atributos de contraste, não como temas. A *skill* não avança enquanto esse registro não existe.

A Fase 1, leitura inicial, corresponde à primeira fase de Braun e Clarke, de familiarização com os dados, e vem antes de qualquer código. O agente inventariou o corpus (13 arquivos: as 12 transcrições mais o índice de sessões), percorreu as doze sessões espalhadas por Belém, Recife e Curitiba nas quatro faixas de renda e produziu anotações por sessão, com o perfil do grupo, surpresas e passagens de interesse, sem gerar um único código.

A Fase 2, códigos, corresponde à segunda fase de Braun e Clarke, de geração de códigos iniciais, ancorou cada código em excertos com localização precisa. Foram 54 códigos, de C01 a C54, cada um com significado, excerto representativo, localizador no corpus e nota interpretativa. A título de exemplo, o C01, "Democracia como procedimento eleitoral e liberdade abstrata", sustentou-se em "democracia é você ter a liberdade de expressão, votar em quem você quer, livre arbítrio" (BEL-0-2, linha 16); o C04, "Autoritarismo como imposição e 'não ter voz'", em "não te dá voz, onde você não tem direitos, escolhas" (BEL-5-10, linha 13); e os C42 e C43, "Programas sociais como necessários com controle" e "Programas sociais como 'acomodação' e manipulação", disputaram o mesmo material da proteção social. A verificação de distribuição confirmou que os códigos cobriam as doze sessões.

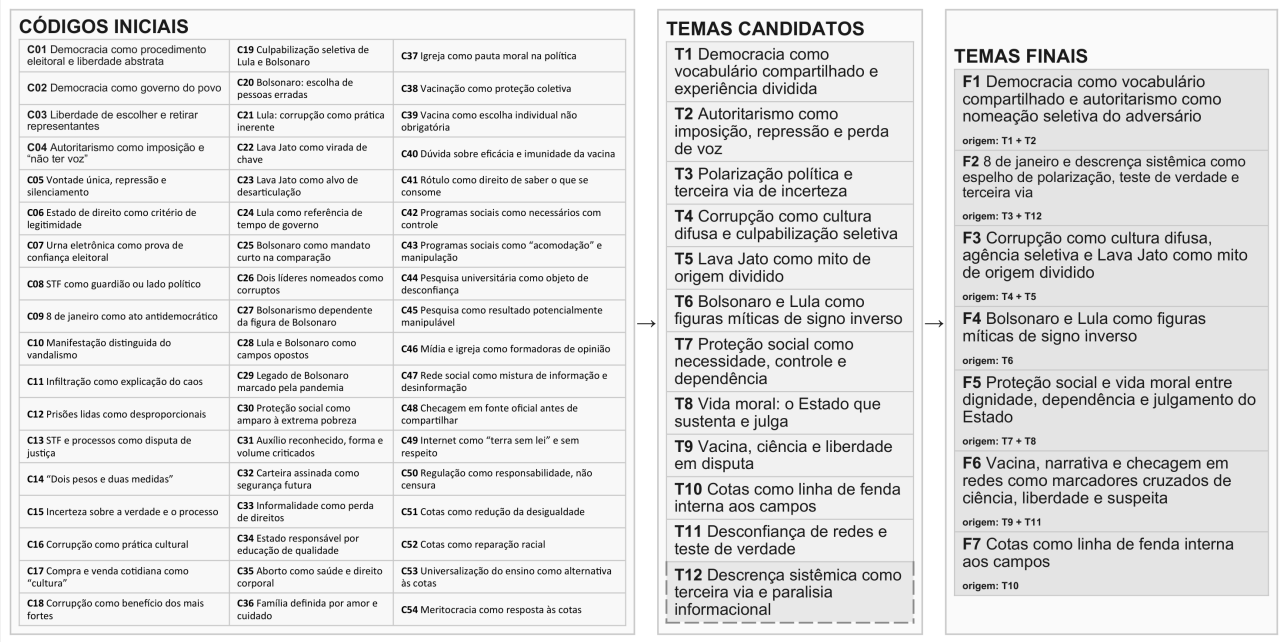
A Fase 3, temas candidatos, corresponde às fases de busca inicial por temas de Braun e Clarke, construiu tais elementos por padrões de significado compartilhado, e não por agrupamento por assunto. Cada tema candidato precisou carregar um conceito organizador e uma afirmação interpretativa, nunca um rótulo de domínio; renda e cidade foram recusadas como tema, porque entram como atravessamento, não como chave. Do mapa de 54 códigos saíram 11 candidatos, de T1 a T11. O primeiro, "Democracia como vocabulário compartilhado e experiência dividida", argumentava que quase todos definem democracia pela mesma fórmula (voto, liberdade, escolha), mas que o sentido prático se bifurca conforme a experiência da derrota eleitoral, da confiança institucional e da renda. O que não se encaixava em nenhum candidato foi para uma pasta provisória de sobras, em vez de ser descartado.

A Fase 4, de contestação, não aceitou o mapa temático sem exame. O agente integrador exigiu uma leitura alternativa plausível e um teste de contradições e sobras, e a leitura rival veio em forma de hipótese: o organizador mestre dos sentidos não é a renda, não é a cidade, mas a posição eleitoral declarada nas sessões pareadas, logo renda e cidade permaneceriam como contexto de produção de experiência material. A rival não derrotou nenhum tema, e essa não era sua função; sua função era forçar o exame. Ela forçou algumas mudanças: o T3 foi revisado para incorporar a terceira via, o "não sei se foi verdade"; o T9 e o T10 foram fortalecidos com casos de dissidência; o T11 foi dividido, porque misturava dois movimentos distintos, o de construir narrativa e o de desconfiar da checagem; e nasceu um tema novo, o T12, "Descrença sistêmica como terceira via e paralisia informacional".

Ter 12 temas mostra amplitude analítica, mas dificuldade de síntese. Por isso, na própria *skill*, estabeleci como regra de ouro mirar entre 4 e 8 temas, se o corpus permitir. A agregação não foi aritmética, pois cada fusão precisou demonstrar que dois temas falavam do mesmo fenômeno. Assim, o T1 e o T2 se fundiram, porque nomear a democracia e nomear o autoritarismo são o mesmo ato de posição no léxico político; o T3 e o T12 se fundiram, porque a descrença sistêmica é a face da mesma polarização que o 8 de janeiro expôs, e a terceira via do T3 era a paralisia do T12 em ato; o T4 e o T5 se fundiram, porque a Lava Jato é o mito de origem da mesma cultura difusa de corrupção que culpabiliza seletivamente; o T7 e o T8 se fundiram, porque proteção social e vida moral são as duas faces da mesma relação, a do Estado que sustenta e julga a mesma população; o T9 e o T11 se fundiram, porque vacina, narrativa e checagem são variações da mesma suspeita epistêmica, e a desconfiança de redes migrou para o tema do teste de

verdade. O T6 e o T10 resistiram à fusão, por saturação e distinção demonstrada. Os doze temas iniciais viraram sete (Figura 1), uma aproximação mais substantiva e mais fácil de recontar a história contida nos temas, como proposto por Braun e Clarke.

Figura 1 – Fluxo da construção dos temas, dos códigos iniciais aos temas candidatos e aos temas finais



Fonte: elaborado pelo autor, 2026.

A Fase 5, conforme o método das autoras, de definição e nomeação, seguiu com os sete temas restantes. Cada tema agora precisou apresentar nome, conceito organizador, afirmação interpretativa, genealogia (dos 12 temas precedentes até os códigos), excertos-chave para interpretação, a leitura rival recusada e a zona de incerteza. Antes disso, os excertos passaram por verificação literal contra o texto original, e o resultado confirmou que as 40 citações candidatas realmente existem no corpus original, porém nenhuma delas foi publicada na forma literal, tendo sido parafraseadas para impedir a identificação dos participantes.

O processo mostra que os sete temas não foram descobertos no corpus; foram construídos, e construídos sob disputa. Os temas finais foram estes: Democracia como vocabulário compartilhado e autoritarismo como nomeação seletiva do adversário; 8 de janeiro e descrença sistêmica como espelho de polarização, teste de verdade e terceira via; Corrupção como cultura difusa, agência seletiva e Lava Jato como mito de origem dividido; Bolsonaro e Lula como figuras míticas de signo inverso; Proteção social e vida moral entre dignidade, dependência e julgamento do Estado; Vacina, narrativa e checagem em redes como marcadores cruzados de ciência, liberdade e suspeita; e Cotas como linha de fenda interna aos campos.

A Fase 6, de relatório, que corresponde à sexta fase de Braun e Clarke, foi a etapa de amarração. O texto contou a história analítica em corpo principal e apêndice auditável. No corpo principal, a pergunta de pesquisa amarra o corpus à narrativa dos sete temas, cada um com excertos-âncora interpretados, variações entre fontes e a rival recusada registrada ao lado da tese que venceu. No apêndice, foram reunidas as fichas completas e contestáveis de cada tema final, contendo sua genealogia, os excertos-âncora verificados, a rival interpretativa, as evidências decisivas, os resíduos, os limites e as incertezas da análise. Também foram incluídos o manifesto de uso de IA, a matriz de delegação por fase, os registros de interação, o mapa humano, a auditoria de cobertura, a verificação dos excertos, os pontos de controle e as decisões adversariais.

Um mecanismo extra para garantir a qualidade da codificação está na verificação que busca cada trecho citado no texto original e separa o que é idêntico do que exige normalização, e somente o idêntico vira aspa. Sem essa conferência, parte intrínseca da *skill*, a IA estaria mais livre para inventar e levar trechos incorretos ou imprecisos à avaliação final. Essa regra vale para a análise feita sobre o corpus original; nesta publicação, os excertos foram parafraseados para impedir a identificação dos participantes.

Avaliação do método

A *skill* foi submetida a um teste comparativo cego com 126 execuções, sessenta e três sem a *skill* e sessenta e três com ela. As vinte e uma situações do teste simulam os atalhos que o protocolo deve recusar, como pular a leitura inicial ou tratar um resumo de assunto como tema. Um agente de IA baseado no modelo GLM 5.2 da Zhipu AI que não participou da geração dos materiais aplicou uma avaliação de 14 critérios a cada resultado (*output*), verificando itens como a presença de excertos literais e a separação entre proposta e decisão. Na avaliação cega, a condição com a *skill* obteve maior aderência aos crité-

rios metodológicos e foi avaliada pelo agente de IA como tendo obtido melhores resultados. O resultado é favorável, sem estabelecer validade, porque nenhum teste substitui a avaliação humana.

DISCUSSÃO

A *skill* reduz a distância entre o *chat* simples e o fluxo de trabalho programático ("agêntico") sem exigir programação do pesquisador. O protocolo da análise temática reflexiva, distribuído em sete fases com verificações, rotas e contestação, vira comportamento observável do agente, e cada fase produz insumos que podem ser examinados, auditados e contestados. O ganho está em tornar o processo reconstruível por terceiros, reforçando princípios de replicabilidade, sem substituir o julgamento humano.

Na minha opinião, o uso de inteligência artificial na pesquisa qualitativa já ocorre em pesquisas, laboratórios e orientações, sem ser enunciado e sem documentação adequada. A *skill* procura responder a esse problema ao organizar critérios, fases e verificações em um artefato único, que qualquer pesquisador pode inspecionar, adaptar e criticar. Um parecerista que discorda de uma análise pode instalar a *skill*, aplicá-la com parâmetros distintos e identificar onde a interpretação diverge; um orientador pode usá-la como objeto de ensino. Na prática, o usuário pode inclusive modificar a *skill* como ele desejar ou mesmo desenhar uma do zero simplesmente pedindo ao agente de IA. Neste teste inicial, optei por fazer uma avaliação apenas por dois agentes de IA, mas criar critérios para avaliar a qualidade por validadores humanos é certamente um caminho a seguir.

A promessa de transparência encontra terreno fértil na pesquisa qualitativa em saúde e enfermagem, campo no qual as exigências de relato transparente crescem[3,13]e a análise temática de dados como entrevistas e grupos focais vem se consolidando na área[14].

É preciso admitir, desde o início, certos limites para tais análises. A inteligência artificial não tem vivência, julgamento situado ou reflexividade humana, e a verificação de citações não garante qualidade interpretativa; ou ainda, o inventário de arquivos gerados pela *skill* confirma a identidade do documento, não o que o agente leu nele, e a concordância entre agentes não equivale a triangulação humana. Custos e assinaturas continuam a criar barreiras, e planos caros pesam mais para pós-graduandos e docentes brasileiros[9]. Dados sensíveis exigem cuidado redobrado, já que falas de participantes processadas por serviços de terceiros impõem a conferência do consentimento original e das regras de proteção de dados pessoais, e o processamento em serviços estrangeiros envolve transferência internacional de dados, que o pesquisador deve avaliar nos termos de cada fornecedor. A decisão sobre o que pode sair do computador pertence ao pesquisador e ao comitê de ética[15].

A rota autônoma, na qual o agente percorre as fases sem intervenção, segue experimental e não declara equivalência com a codificação. Ela torna visível o que acontece quando a máquina executa o protocolo completo e produz material que outro pesquisador pode auditar, contestar ou usar como leitura inicial ou mesmo como uma leitura rival.

CONCLUSÃO

Este texto buscou apresentar uma forma bastante simples de rodar uma análise temática automatizada, bastando instalar o programa de agente, baixar a *skill*, abrir a pasta com o material a ser analisado e pedir a análise em linguagem natural. Este caminho me parece inovador e promissor, uma vez que ele facilita o acesso de qualquer pesquisador sem exigir conhecimentos de programação, ao mesmo tempo que gera materiais para auditar, questionar ou mesmo embasar a análise. Quem colava trechos no *chatbot* agora pode pedir a análise inteira em linguagem natural e mostrar, documento por documento, de onde veio cada tema e cada fala.

A demonstração em corpus real mostrou que o fluxo completo é possível, com cobertura total e cada citação conferida contra o original. O teste cego é compatível com a redução de desvios metodológicos em relação ao mesmo modelo sem orientação, e a portabilidade mostrou que o pacote não depende de um único fornecedor. Nenhuma dessas evidências autoriza afirmar equivalência humana ou validade automática; elas sustentam viabilidade, transparência e auditabilidade, propriedades que a pesquisa qualitativa precisa defender na era da inteligência artificial.

O limite óbvio neste estágio é a falta de validação com um ou mais pesquisadores humanos e com análises temáticas de qualidade, algo que será realizado nas próximas etapas da pesquisa. A validação por dois agentes distintos de IA também tem limites, pois pode apenas evidenciar que essas tecnologias apresentam vieses similares, no entanto, diversos experimentos pessoais e o experimento de validação indicam melhor desempenho na codificação de aspectos qualitativos.

O valor científico de um agente qualitativo está no quanto ele permite a outros reconstruir, contestar e transformar o que produziu. A interpretação final e a responsabilidade por cada afirmação continuam humanos. A ferramenta deve ajudar o pesquisador a pensar melhor, nunca a pensar no lugar dele.

CONFLITO DE INTERESSES

O autor declara não haver conflito de interesses.

FINANCIAMENTO

O presente trabalho foi realizado com apoio de bolsa de produtividade do CNPq.

USO DE INTELIGÊNCIA ARTIFICIAL

A *skill* descrita neste artigo foi desenvolvida e avaliada com o uso de ferramentas de inteligência artificial generativa. O ChatGPT Luna 6 foi usado para os testes da *skill* de análise temática e também para correções e melhorias no texto, além de ter auxiliado na tradução. A redação deste manuscrito contou com assistência dessas ferramentas, com revisão e edição final pelo autor, que assume integralmente a responsabilidade pelo conteúdo.

REFERÊNCIAS

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]. Disponível em: 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. No prelo.
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. Disponível em: 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]. Disponível em: https://www.portcom.intercom.org.br/ebooks/arquivos/livro-diretrizes-ia.pdf

CONTRIBUIÇÃO DE AUTORIA

Concepção do estudo: Sampaio RC.
Obtenção de dados: Sampaio RC.
Análise de dados: Sampaio RC.
Interpretação dos dados: Sampaio RC.
O autor se responsabiliza pela redação textual e revisão crítica do conteúdo intelectual, pela versão final publicada e por todos os aspectos éticos, legais e científicos relacionados à exatidão e à integridade do estudo.

© 2026 Autores. Os autores mantêm os direitos autorais e concedem à Online Brazilian Journal of Nursing o direito de primeira publicação. Este artigo está licenciado sob a licença Creative Commons Atribuição 4.0 Internacional (CC BY 4.0).