26º ENANCIB
Encontro Nacional de Pesquisa e Pós-Graduação em Ciência da Informação
A Amazônia na sociedade global em rede: assimetrias e potencialidades informacionais, tecnológicas e ambientais
Belém, Pará, 09 a 13 de novembro de 2026
Universidade Federal do Pará

XXVI ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO - XXVI ENANCIB

2177-3688

GT 5 - Política e Economia da Informação

O QUE O GROK SABE SOBRE SAÚDE? ANÁLISE DE POSTAGENS NO X

WHAT GROK KNOWS ABOUT HEALTH? *ANALYSIS OF POSTS ON X*

Frederico Oliveira - Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)
Larissa de Araújo Alves - Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)
Graziela Barros Gomes - Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)
Cristiane Henrique - Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)
Tiago Emmanuel Nunes Braga - Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)

Modalidade: Trabalho Completo

Resumo: Analisa a mediação de ferramentas de inteligência artificial generativa na circulação da informação científica, tomando como objeto as respostas do Grok sobre saúde e vacinação no X. Busca compreender como a IA responde a usuários em debates sobre saúde e que tipos de evidências são mobilizadas em suas respostas. Para isso, foram analisadas postagens coletadas entre junho de 2025 e maio de 2026 a partir da base do projeto Painel On-line de Detecção de Narrativas Antivacina. O estudo adota abordagem quanti-qualitativa, combinando análise textual automatizada, identificação de fontes por expressões regulares, agrupamento semântico por embeddings SBERT e clusterização HDBSCAN, além de análise qualitativa de uma amostra de postagens sobre dengue. Os resultados indicam que o Grok tende a apresentar respostas com aparência de neutralidade e segurança, frequentemente mobilizando fontes oficiais, científicas, jornalísticas e genéricas. Contudo, observa-se baixa transparência na indicação das fontes consultadas, ausência de links diretos e pouca hierarquização entre diferentes tipos de evidência. Conclui-se que, embora o Grok contribua para sumarizar informações disponíveis on-line, sua atuação pode produzir uma informação apenas aparentemente científica, ao equiparar fontes de naturezas distintas e não explicitar adequadamente os caminhos metodológicos que sustentam suas respostas.

Palavras-chave: Inteligência artificial generativa; informação científica; Grok; saúde; plataformas digitais.

Abstract: This article analyzes the mediation of generative artificial intelligence tools in the circulation of scientific information, focusing on Grok's responses about health and vaccination on X. The study seeks to understand how the AI responds to users in online health debates and what types of evidence are mobilized in its answers. To this end, posts published between June 2025 and May 2026 were analyzed using data from the Online Panel for the Detection of Anti-Vaccine Narratives project. The research adopts a quantitative-qualitative

approach, combining automated textual analysis, source identification through regular expressions, semantic grouping using SBERT embeddings and HDBSCAN clustering, as well as a qualitative analysis of a sample of posts about dengue. The results indicate that Grok tends to provide answers that convey neutrality and certainty, frequently mobilizing official, scientific, journalistic, and generic sources. However, the analysis reveals low transparency in the indication of consulted sources, a lack of direct links, and limited hierarchization among different types of evidence. The study concludes that, although Grok contributes to summarizing information available online, its performance may produce only apparently scientific information, as it equates sources of different natures and does not adequately explain the methodological paths that support its responses.

Keywords: Generative artificial intelligence; scientific information; Grok; health; digital platforms.

1 INTRODUÇÃO

Nas plataformas digitais, a disseminação da informação científica envolve *chatbots*, interfaces, redes de distribuição de conteúdo (CDN), dentre outros recursos computacionais e humanos. A partir desta perspectiva, busca-se compreender a mediação de ferramentas de inteligência artificial generativa (IAG) na circulação da informação científica. Para tanto, é realizada uma análise de publicações do Grok na discussão on-line sobre saúde e vacinação no X. Tal escolha metodológica implica na compreensão de como as perguntas dos usuários da plataforma são respondidas pela ferramenta de IAG.

Com a introdução do Grok no X, as práticas de consumo da informação se alteraram. O *bot*[1] é utilizado como árbitro da veracidade de uma informação e recurso para a sumarização de conteúdo. Por estar integrado à plataforma, o Grok interfere diretamente na conversação on-line no X, sendo convocado por usuários para explicar, contextualizar ou contestar publicações, assim como para apontar quem estaria correto em debates. Tal dinâmica acompanha mudanças no perfil de busca da informação, tal como o uso crescente do TikTok (Adobe, 2026) e da IAG (Fundação Itaú; Datafolha, 2025) como ferramentas de busca, indicando a disponibilidade e uso de novas ferramentas e plataformas. A partir de tal compreensão, e da premissa de que a produção, a disseminação e o consumo da informação científica na contemporaneidade são atravessados pelas materialidades das plataformas

[1] É um assistente de inteligência artificial generativa (IAG), desenvolvido pela xAI e lançado em novembro de 2023, com integração direta à plataforma X. O Grok pode recorrer a postagens públicas do X e realizar buscas na web em tempo real para responder às consultas dos usuários. Em 2024, foram incorporados recursos de busca na web e de citação de fontes. Além de poder ser acessado por uma interface conversacional própria, o Grok está integrado ao X, onde pode responder a solicitações dos usuários, como pedidos de explicação, resumo ou verificação de conteúdos. O Grok esteve envolvido em controvérsias relacionadas a supostos vieses políticos, além da presença de incorreções nas respostas aos usuários.

digitais, investigar como o Grok produz informação sobre saúde mostra-se relevante para se aprofundar sobre essas transformações.

Duas perguntas guiam este estudo: De que modo o Grok responde aos usuários que perguntam questões sobre saúde? Que tipo de evidências o Grok apresenta em suas respostas? Acredita-se que as respostas são fundamentadas em fontes genéricas, sem indicação direta de documentos ou hierarquização. Em relação ao tom da resposta, espera-se que o Grok adote um teor mais informativo, aparentando neutralidade e polidez, prática comum adotada pelas IAG (Paiva, 2025). Para responder a tais questões, são feitas análises automatizadas e manuais de postagens[2] do *bot* sobre saúde e vacinas entre junho de 2025 e maio de 2026.

A primeira questão envolve a identificação de padrões discursivos, assim como os principais temas a partir dos quais as interações se organizam. Para a análise das evidências apresentadas pelo assistente, há identificação automática de menções a fontes e marcadores de atribuição, complementada com análise qualitativa. O artigo ainda apresenta uma reflexão sobre a plataformização da informação científica e a mediação de ferramentas de IAG.

2 PLATAFORMAS E INFORMAÇÃO CIENTÍFICA

O fato científico, como demonstra Latour (1994), é resultado das condições específicas de sua produção. A partir de instrumentos de acesso específicos, são produzidas cadeias de referência que permitem ir do objeto ao relato, e vice-versa. Cada instrumento, contudo, produz um resultado distinto. A qualidade da ciência que se produz está relacionada à capacidade de garantir o deslocamento do relato ao objeto, e vice-versa, a partir das provas produzidas por meio de instrumentos de acesso (Latour, 2019).

É possível extrapolar tal compreensão para a disseminação da informação científica. Além de ser necessário garantir o encadeamento correto de referências para que, de fato, seja possível falar bem[3] , no relato científico, é preciso considerar as intervenções dos meios

[2] A integração do Grok ao X torna difícil classificar as respostas analisadas. A estratégia de coleta adotada nesta pesquisa implica na coleta dessas respostas como publicações de autoria do Grok, ainda que produzidas em resposta a uma interação do usuário e vinculadas a tal postagem. Interessa, assim, observar a dupla natureza dessas respostas: o Grok é apresentado ao mesmo tempo pela plataforma como sistema autônomo, enquanto também opera como uma interface do X. Entende-se, contudo, ser importante considerar que a mediação da informação promovida pelo Grok é resultado de um entrelaçamento entre usuário, interface, algoritmos, a arquitetura informacional do assistente, dentre vários outros objetos - a mediação está entre estes atores, sendo atravessada por cada um deles.

[3] Falar bem, para Latour (2019), não é apenas descrever os atores de cada rede, mas entender o que há de específico naquela rede. Em outras palavras: para entender as redes da Ciência não posso usar os recursos de

usados para a distribuição de tais conteúdos. Nas plataformas digitais, por exemplo, imagens perdem resolução, metadados são apagados, conteúdos são classificados automaticamente, são ranqueados por algoritmos de relevância pública, dentre outras intervenções. Pensar a disseminação da informação científica em um contexto de plataformização (Poell; Nieborg; Van Djick, 2020), portanto, demanda a compreensão das condições de produção, de circulação e de consumo mediado pelas plataformas.

Novos fatores entram em cena: a otimização para resultados de busca (Barcelos; Oliveira; Souza, 2025), a adesão a determinadas pesquisas - ainda que incipientes - por interesse político (Cesarino, 2022), a circulação de conteúdos entre plataformas (Machado *et al*., 2020), o pagamento de influenciadores digitais para disseminar um tratamento não aprovado (Affonso, 2026), dentre outros. Mais recentemente, a introdução de ferramentas de IAG implicou na sumarização e simplificação da informação científica disponível em periódicos científicos abertos, por exemplo. No entanto, tais recursos consultam documentos distintos, sem necessariamente hierarquizar as fontes consultadas por sua qualidade. Surgem, então, questionamentos sobre a confiabilidade da informação das IAGs, especialmente em temas como a saúde (Agar; Key, 2026).

O uso das ferramentas de IAG na disseminação da informação científica implica em alguns desafios metodológicos, éticos e técnicos (Oliveira; Barcelos; Campos, 2025). Uma das questões mais relevantes é a análise da precisão e da acurácia das informações por elas resumidas. Por se tratar de uma lógica estatística e preditiva (Campolo; Crawford, 2020), e considerando a atualização em tempo real de recursos como o Grok e o ChatGPT, não há garantia de que as mesmas fontes sejam sempre consultadas, tampouco que o mesmo texto será apresentado a distintos usuários. Assim, é difícil mensurar o grau de precisão da informação apresentada, haja vista que a resposta é estatisticamente provável para o momento específico da pergunta. Há, portanto, o desafio metodológico de analisar a confiabilidade, a precisão e a acurácia de questionamentos a tais recursos, especialmente considerando múltiplos usuários.

Uma estratégia possível é o desenvolvimento de cadernos de campo (Lemos, 2024, Oliveira; Barcelos; Campos, 2025), assim como se pode consultar relatórios de *benchmarking* das próprias empresas que apontam a frequência com que tais recursos alucinam (OpenAI,

outras áreas, como o Direito. Falar bem, neste sentido, é entender de onde se fala e como o conhecimento é construído neste contexto. Para falar sobre a Ciência, é preciso falar a partir de seus recursos e sua forma de produzir / evocar o mundo.

2025). Trata-se, contudo, de uma análise situada, cuja generalização é limitada. Para além disso, observar a quantidade de erros e acertos, por si só, não permite compreender como as IAG fundamentam suas considerações. Independentemente da metodologia, contudo, há limites nas análises, já que, em toda pesquisa em IAG, as condições de observação condicionam os resultados.

3 PROCEDIMENTOS METODOLÓGICOS

Trata-se de pesquisa de caráter descritivo e exploratório, com abordagem quanti-qualitativa. A caracterização de como o Grok responde aos usuários combina análise dos padrões discursivos e agrupamento semântico das postagens; já a investigação das evidências mobilizadas considera a identificação das fontes mencionadas e de sua forma de apresentação. A análise quantitativa do corpus completo é posteriormente complementada por uma análise qualitativa de respostas sobre a dengue.

3.1 COLETA DE DADOS

Publicações em redes sociais sobre saúde e vacinação são coletadas diariamente, desde junho de 2024, pela pesquisa *Painel On-line de Detecção de Narrativas Antivacina*[4]. Considerando tal base, em 20 de maio de 2026, foram filtradas todas as postagens do X, correspondente a 60,8% do total de publicações armazenadas na base, resultando em 1.112.525 publicações. Uma nova seleção foi aplicada, considerando exclusivamente postagens de autoria do Grok. Na ocasião da coleta, o Grok era responsável por 32.291 postagens (2,9% das publicações no X). A coleta considerou publicações entre 14 de junho de 2025 e 11 de maio de 2026.

No que tange à coleta de dados, vale ressaltar que existem limitações relacionadas à natureza do banco de dados utilizado e à própria interface de programação de aplicações (API) do X. Na coleta desenvolvida para a base, foram recuperados em sua completude *posts* com

[4] Com mais de 1,8 milhão de postagens, a base reúne postagens com foco nas temáticas de vacinas e vacinação e é atualizada diariamente a partir de consultas a API de plataformas e também por meio de ferramentas complementares de raspagem de dados de postagens públicas. A primeira coleta foi realizada em junho de 2024. Desde então, novos termos de busca foram inseridos a fim de compreender melhor a discussão on-line sobre saúde. Na ocasião da coleta realizada para este trabalho, a base era composta por postagens no X (60,8%), no Instagram (17,46%), no Facebook (9,72%), no TikTok (6,94%) e no YouTube (4,87%). Vale ressaltar, no entanto, que toda coleta em redes sociais é contingente às limitações das APIs e de raspagem, ao contexto de publicação (postagens podem ser excluídas, por exemplo) e aos termos de uso utilizados.

até 280 caracteres[5], sendo que textos maiores não foram recuperados em sua totalidade. Ademais, a coleta em uma base de dados já existente limita a recuperação de novas postagens. Por outro lado, é uma base cuja atualização é diária e com estratégias de coleta consolidadas, que permite estudos longitudinais. Considerando que o objetivo deste estudo é compreender o comportamento do Grok perante questionamentos dos usuários sobre saúde, tais limitações não trazem prejuízos significativos.

3.2 ANÁLISE TEXTUAL

Considerando o volume de postagens, optou-se por estratégias automatizadas de análise textual das mensagens de autoria do Grok, realizada em duas etapas. Na primeira, de caráter automatizado, foram analisadas as postagens do Grok presentes no corpus, após normalização textual. Expressões regulares foram utilizadas para identificar menções a fontes e padrões lexicais associados a diferentes formas de resposta da ferramenta. Paralelamente, *embeddings* SBERT e clusterização HDBSCAN foram empregados para mapear os principais agrupamentos temáticos do *corpus*; bigramas e trigramas auxiliaram na interpretação desses agrupamentos. Essa etapa não teve como objetivo classificar qualitativamente cada resposta, mas identificar padrões gerais e orientar a seleção do recorte para análise manual. O Quadro 1 apresenta o fluxo de análise:

Quadro 1 – Fluxo da análise quantitativa

Etapa 1 - Todas as respostas do Grok
Normalização textual: minúsculas, Unicode, remoção de acentos e caracteres especiais, padronização dos espaços e limpeza de menções a perfis.
Identificação de fontes: uso de expressões regulares (*regex*) para identificar *links* e menções a fontes.
Identificação de fontes específicas: também por meio de *regex*, foram coletadas menções a fontes específicas.
Grau de certeza: identificação de padrões de resposta do Grok e uso de palavras que denotam dúvida ou surpresa.
Análise semântica automatizada: *embedding* SBERT e clusterização por meio de HDBSCAN, para identificar *clusters* semânticos.
Identificação de bigramas e trigramas: uso de TF-IDF para filtragem de palavras e posterior

[5] Trata-se de limitação da ferramenta utilizada para a coleta de publicações no X.

Etapa 1 - Todas as respostas do Grok
identificação de bigramas e trigramas para, assim, interpretar os *clusters*.

Fonte: elaboração própria (2026).

Para tanto, fez-se necessária a preparação do texto: as postagens foram convertidas em minúsculas, ajustou-se sua codificação, assim como foram removidas menções a usuários feitas pelo Grok. Foi desenvolvido, então, um dicionário de padrões lexicais para identificar como o Grok responde aos usuários (Quadro 2). Foram identificados *links* compartilhados, menções a veículos jornalísticos, órgãos governamentais, organismos internacionais, documentos, dentre outros. Também fez-se a identificação de menções a fontes específicas, selecionadas previamente pelos pesquisadores a partir da leitura de amostra do *corpus*.

Quadro 2 – Expressões consideradas

Tipo	Expressões*
Mídia jornalística brasileira	“G1”, “Globo”, "O Globo”, “Folha de S. Paulo”, “Estadão”, “UOL”, “Metrópoles”, “CNN Brasil”, “Agência Brasil”, “Poder 360”, “Gazeta do Povo”, “Valor Econômico”, “Veja”, “Terra”, “R7”, “Jovem Pan”, “Correio Braziliense”, “Exame”, “Nexo”, “Carta Capital”, “Brasil de Fato”, “O Antagonista”, “Infomoney”
Mídia jornalística internacional	“BBC”, “CNN”, “Reuters”, “New York Times”, “The Guardian”, “The Washington Post”, Associated Press”, “AFP”, “Financial Times”, “The Economist”, “The Intercept”
Agências de checagem	“Aos Fatos”, “Agência Lupa”, “Fato ou Fake”, “Boatos.org”, “E-farsas”, Estadão Verifica”, “Projeto Comprova”, “verificadores independentes”
Órgão público e regulador brasileiro	“Anvisa”, “Ministério da Saúde”, “SUS”, “Governo Federal”, “Secretaria de Saúde”, “STF”, “STJ”, “TSE”, “TCU”, “CGU”, “Ministério Público”, “MPF”, “PGR”, “Polícia Federal’, “Receita Federal”, “Banco Central”, “CVM”, “IBGE”, Ibama”, “Inpe”, “Congresso Nacional”, “Senado”, “Câmara dos Deputados”, “CPI da Covid”
Organismos internacionais e multilaterais	“OMS”, “OPAS”, “ONU”, “Unicef”, “Unesco”, “Banco Mundial”, “FMI”, “OCDE”, “Human Rights Watch”
Órgãos técnicos internacionais	“CDC”, “FDA”, “NIH”, “EMA”, “NHS”, “Health Canada”, “NASA”, “NOAA”
Instituição científica	“UFRJ”, “Universidade Federal do Rio de Janeiro”, “Instituto de Ciências Biomédicas”, “Fiocruz”, “Butantan”, “USP”, “Unicamp”, “UFMG”, “Harvard”, “Oxford”, “Stanford”, “MIT”, “Mayo Clinic”, “Lowy Institute”,

Tipo	Expressões*
	“Cristália”, “pesquisadores / cientistas”.
Revistas, bases e publicações científicas	“PubMed”, “SciELO”, “Nature”, “Science”, “The Lancet”, “New England Journal of Medicine”, “JAMA”, “BMJ”, “DOI”, “ensaio clínico”, “estudo científico”
Bases de dados, painéis e repositórios	“Our World in Data”, “Wikipedia”, “Wikidata”, “Google Scholar”, “Worldometer”
Pesquisa de opinião	“Datafolha”, “Quaest”, “Poder Data”, “Atlas Intel”, “Data Senado”, “Paraná Pesquisas”
Documento	“Relatório”, “Boletim epidemiológico”, “nota técnica, “parecer técnico”, “dados oficiais”, “base de dados”, “levantamento”, “estatísticas”, “lei”, “decreto”, “portaria”, “resolução”, “decisão judicial”, “processo judicial”, “inquérito”, “depoimento”
Fonte genérica	“Fontes”, “fontes oficiais”, “fontes independentes”, “relatos”, “relatos confirmados”, ‘dados recentes”, “informações recentes”, “segundo fontes”, “segundo relatos”, “de acordo com fontes” com base em relatos”, “com base em informações”
Marcador de atribuição	“Segundo”, “de acordo com”, “conforme”, “com base em”, “reportagem de”, “noticiado por”, “divulgado por”, “publicado por”.

* Destaca-se que algumas expressões regulares utilizadas podem resultar em falsos positivos, de modo que se considera que os resultados da etapa quantitativa sejam exploratórios, a fim de melhor compreender a estrutura de respostas do Grok.

Fonte: elaboração própria (2026).

Em etapa complementar, foi desenvolvida uma análise semântica por meio de *embeddings* SBERT, com posterior clusterização a partir de HDBSCAN. Para apoiar a interpretação de cada um dos *clusters*, foram identificados conjuntos de duas (bigramas) ou três (trigramas) palavras mais relevantes. Deste modo, foi possível identificar os principais assuntos discutidos pelo Grok. A partir da compreensão dos temas mais discutidos no *corpus*, foi selecionada amostra de postagens para análise qualitativa.

Em relação à análise quantitativa, pode-se apontar como limitações a estratégia de identificação de fontes por meio de expressões regulares. Embora o dicionário de expressões tenha sido desenvolvido após análise exploratória dos dados, é possível que termos específicos não tenham sido considerados, enquanto outros sejam super-representados.

Além disso, quando se busca identificar fontes citadas a partir de termos como “segundo”, falsos positivos são possíveis. As expressões regulares identificam a grafia, não a

classe gramatical, de modo que se identifica “segundo” tanto como preposição (desejável) como adjetivo, advérbio, conjunção, numeral e substantivo (indesejáveis). Outra possibilidade: “ministério da saúde” pode se referir a “Ministério da Saúde” brasileiro ou a “Ministério da Saúde de Gaza”, por exemplo. É possível, portanto, que menções a instituições diferentes sejam consideradas conjuntamente, assim como algumas fontes citadas não serem identificadas. De tal modo, a análise qualitativa busca reduzir tais lacunas e equívocos de classificação.

Ressalta-se, contudo, que o artigo aqui apresentado busca compreender como o Grok responde aos usuários. As limitações destacadas, embora interfiram nos resultados quantitativos apresentados, não impedem a compreensão dos perfis de resposta da IA. É possível, a partir da estratégia de análise adotada, a compreensão geral de como o Grok intervém no debate sobre saúde. A existência de uma etapa complementar, com análise qualitativa, ajuda a reduzir as limitações supracitadas.

3.3 ANÁLISE QUALITATIVA

A partir do agrupamento semântico, foram identificados os temas mais relevantes nas publicações do Grok (Quadro 3). Compreendeu-se que a dengue foi o tema mais discutido, com 1.098 mensagens. Foi realizada uma filtragem das mensagens com mais de uma (1) curtida, resultando em 151 postagens, excluídas duplicatas[6]. Tal opção se justifica pois se buscou identificar publicações do Grok que receberam endosso de mais de um perfil - não apenas daquele que fez a pergunta à IA. Considerando a limitação da coleta, fez-se a busca manual pelas postagens, a fim destas serem analisadas em sua completude. Algumas publicações haviam sido apagadas, não sendo possível sua recuperação. Ressalta-se que os critérios de seleção da amostra qualitativa não foram empregados na análise automatizada do corpus completo, mas apenas na definição do subconjunto submetido à codificação manual.

Quando um usuário apaga uma postagem em que fez uma pergunta ao Grok, a resposta da IA também fica indisponível. Tal comportamento é interessante por duas questões: a exclusão de perguntas ao Grok pode indicar que um usuário, quando corrigido pela IA, sentiu-se constrangido e apagou o conteúdo; a exclusão da resposta do Grok pode

[6] Observou-se a presença de postagens distintas com o mesmo texto. Em alguns casos, o Grok publica duas vezes (há diferentes IDs).

indicar que publicações do *bot* não são tratadas como postagens de usuários. Afinal, quando um usuário exclui uma postagem, respostas a essa publicação feitas por outros usuários não são apagadas, ainda que a postagem inicial tenha sido retirada. No entanto, no caso de exclusão de uma publicação respondida pelo Grok, a resposta da IA também é apagada. Tal comportamento reforça a ambígua natureza das postagens do assistente, destacando a importância de entender que a mediação da informação pelas IAG é situada, contingente ao contexto e resultado da interação entre humanos, algoritmos, interfaces, plataformas e outros elementos.

A nova amostra foi analisada manualmente, para se identificar: *a)* fonte citada; e *b)* como a fonte foi apresentada. Cada mensagem foi analisada por mais de um codificador[7], a fim de se garantir a confiabilidade da análise. Dentre as limitações desta etapa, cabe apontar que a restrição de coleta automática de *posts* com até 280 caracteres demandou outra coleta, manual, da amostra selecionada. No entanto, quando o usuário apaga a postagem em que pergunta ao Grok, as respostas do *bot* também ficam indisponíveis, o que impediu a recuperação completa das publicações. Por outro lado, tal limitação pode indicar que o usuário tende a apagar postagens com incorreções quando é corrigido pelo Grok.

4 RESULTADOS E DISCUSSÃO

São apresentados, a seguir, os resultados da análise automatizada e, também, da identificação manual de fontes em amostra:

4.1 AS EVIDÊNCIAS DO GROK

Os dez principais agrupamentos semânticos são apresentados no Quadro 3. Observa-se que o debate sobre vacinas e saúde pública é atravessado por temáticas políticas, como a gestão da pandemia. Considerando que as postagens analisadas são respostas do Grok, percebe-se a intervenção da IA na discussão on-line sobre tais temas.

[7] A codificação foi realizada independentemente por dois avaliadores. Para as variáveis binárias, a confiabilidade *intercoder* foi estimada pelo Kappa de Cohen, obtendo-se $\kappa = 0{,}761$ para presença de fonte, $\kappa = 0{,}748$ para identificação do nome da fonte e $\kappa = 0{,}744$ para identificação de links. As divergências foram posteriormente discutidas e resolvidas por consenso.

Quadro 3 - Principais agrupamentos semânticos

Tema	Termos identificados	Publicações
Dengue	dengue; mortes dengue; mortes; wolbachia; mosquitos; contra dengue; aedes; aedes aegypti; aegypti; ministério saúde; ministério; qdenga; saúde; covid; cerca	1098
Miocardite causada por vacinas mRNA	miocardite; jovens; mrna; vacinas; covid; vacinas mrna; raros; jovens sexo; jovens sexo masculino; sexo masculino; miocardite pericardite; risco; pericardite; masculino; sexo	475
Primeira vacina contra a Covid-19	primeira; contra covid mundo; dezembro reino unido; dezembro reino; reino unido; unido; reino; covid mundo; aplicada dezembro; primeira vacina contra; pfizer biontech; aplicada dezembro reino; contra covid; dezembro; biontech	473
Surtos de gripe	gripe; influenza; contra gripe; vírus; aviária; gripe aviária; risco; cdc; febre; vacinas contra gripe; contra; sintomas; vírus influenza; vacina; casos	390
Validade científica das vacinas	vacinas; cdc; estudos; sarampo; imunidade; riscos; vidas; pólio; doenças; oms; raros; mortes; vacina; varíola; milhões vidas	324
Debate político sobre vacina	vacinas; fatos; ciência; evidências; estudos; desinformação; narrativas; saúde; covid; mbl; posso; saúde pública; pública; liberdade; ex	292
Gestão da pandemia de Covid-19	bolsonaro; pandemia; mortes; vacinas; minimizou; gripezinha; mil mortes; tratamentos; mil; tratamentos ineficazes; promoveu; ineficazes; bolsonaro minimizou; hidroxicloroquina; vírus	270
Aprovação emergencial de vacina contra Covid-19 pela Anvisa	primeiras vacinas contra; emergencial primeiras vacinas; uso emergencial primeiras; contra covid coronavac; emergencial primeiras; covid coronavac; uso emergencial; primeiras vacinas; covid coronavac astrazeneca; aprovou uso emergencial; aprovou uso; anvisa aprovou uso; emergencial; coronavac; astrazeneca janeiro	266
Eficácia da vacina contra Covid-19	contra; covid; contra covid; vacinas contra covid; vacinas; vacinas contra; cdc; hospitalizações; estudos; eficácia; mortes; graves; hospitalizações mortes; raros; casos graves	264
Economia brasileira durante a	pib; desemprego; inflação; bolsonaro; lula; covid; recessão; crescimento; world bank; bank; médio; pib médio; dívida; caiu; imf	261

Tema	Termos identificados	Publicações
Covid-19		

Fonte: elaboração própria (2026).

Em 23.184 postagens do Grok foi possível identificar fontes indicadas diretamente, enquanto em 5.610 não há indicação de fontes. Em 860 postagens são citadas fontes genéricas (“estudos”, “levantamento” etc.). No que tange às publicações com indicação de fontes, a partir dos critérios previamente apresentados no Quadro 2, observa-se priorização de fontes oficiais (Quadro 4). Ressalta-se que uma mesma publicação do Grok pode nomear referências distintas, de modo que o somatório das fontes excede o número de postagens.

Quadro 4 - Fontes apresentadas pelo Grok

Apresentação da fonte	n
Fonte oficial	13.609
Fonte genérica	6.930
Fonte científica / Estudo	6.004
Marcador de atribuição	5.056
Documento ou dado oficial	3.803
Organização internacional	3.515
Imprensa nacional	3.330
Órgão técnico internacional	2.617
Imprensa internacional	2.434
Bases de dados	831
Agência de checagem	251
Pesquisa de opinião	67

Fonte: elaboração própria (2026).

A menção a fontes oficiais pode implicar em uma abordagem declaratória, marcada pela mera reprodução do discurso oficial. Embora a informação apresentada pelas fontes oficiais seja, geralmente, tida como verdadeira (Lage, 2008), não se pode olvidar o interesse das fontes (Schmidt, 2011) e eventual falseamento dos dados (Lage, 2008).

Em 5.610 postagens não são mencionadas fontes. A ausência de fontes, contudo, pode representar outras formas pelas quais o Grok busca a informação solicitada pelo usuário. É o que acontece, por exemplo, quando alguém pergunta sobre o teor de um vídeo. Nesses casos, a IA transcreve e sumariza o conteúdo, apresentando ao usuário uma breve descrição. Importa

destacar, contudo, que o Grok não torna tais estratégias transparentes ao usuário, que somente as pode inferir.

Por fim, o Quadro 5 busca sintetizar o perfil discursivo do Grok. A partir da identificação de termos específicos, analisou-se o tipo de ação da IA e as fontes consultadas. Observa-se que a maior parte das postagens busca apresentar fontes que confirmam o que a IA afirma.

Quadro 5 - Perfil discursivo do Grok

Padrão de ação	Termos	Frequência	Fontes*	Posts
Mobilização de evidência	Segundo; de acordo com; conforme; com base em; dados; fontes; relatório; checagem; estudo; pesquisa; artigo; ensaio; URLs.	12041	Não científica	11.246
			Sem fonte explícita	543
			Fonte científica	252
Qualificação / nuance	Não exatamente; não necessariamente; parcialmente; em parte; depende; mas; porém; contudo; no entanto; embora; uma prática; oficialmente; tecnicamente; hiperbólico; não é automático.	8594	Não científica	6.474
			Sem fonte explícita	1.985
			Fonte científica	135
Outra		2283	Sem fonte explícita	1552
			Não científica	681
			Fonte científica	50
Validação / ratificação	Sim, é verdade; Sim, procede; correto; exatamente; é verdade; é verdadeiro; é real; procede; correto; de fato; você tem razão.	2011	Não científica	1531
			Sem fonte explícita	424
			Fonte científica	56
	Trata-se; o caso; o vídeo; a imagem; a notícia; a lei; o projeto; a decisão;	1492	Sem fonte explícita	831

Padrão de ação	Termos	Frequência	Fontes*	Posts
Explicação / contextualização	foi criado; foi aprovado; ocorreu em; desde; no Brasil; nos EUA.		Não científica	625
			Fonte científica	36
Refutação / desmentido**	Não é verdade; Não, é verdade; é mentira; é falso, falso; mentira, incorreto; não procede; sem provas; sem evidências; não há evidências; não há registros.	1387	Não científica	1.059
			Sem fonte explícita	301
			Fonte científica	27
Síntese / listagem	Resumo; em resumo; resumindo; em suma; pontos principais; eis; segue; principais pontos; primeiro; segundo; terceiro.	1247	Sem fonte explícita	610
			Não científica	609
			Fonte científica	28
Cálculo / estimativa quantitativa	Cerca de; aproximadamente; estimado; estima-se; cálculo; percentual; porcentagem; %; milhões, bilhões; R$; por ano; por mês; taxa.	1229	Sem fonte explícita	606
			Não científica	592
			Fonte científica	31
Metadiscurso / defesa de si próprio	Como IA; como uma IA; não sou; não tenho opinião; não tenho viés; imparcial; militante; viés; vamos aos fatos; posso ajudar; quer que eu.	788	Não científica	599
			Sem fonte explícita	179
			Fonte científica	10
Comparação / diferenciação	Compar; diferenc; difere; enquanto; por outro lado; ambos; maior que; menor que; mais do que; menos do qu; entre ** e **.	416	Sem fonte explícita	285
			Não científica	126
			Fonte científica	5

Padrão de ação	Termos	Frequência	Fontes*	Posts
Definição / identificação	O que é; quem é; significa; é um; se refere a; consiste em; chama-se	324	Sem fonte explícita	208
			Não científica	104
			Fonte científica	12
Orientação / recomendação	Procure; consulte; recomendo; evite; verifique; confira; fique atento; fica ligado; cuidado; lave as mãos; se tiver sintomas; médico; especialistas.	266	Sem fonte explícita	150
			Não científica	115
			Fonte científica	1
Acolhimento / interação conversacional	Olá; oi; bom dia; boa tarde; boa noite; de nada; obrigado; claro; beleza; entendo; entendo sua preocupação; entendo sua frustração; entendo seu ponto.	225	Sem fonte explícita	141
			Não científica	81
			Fonte científica	9
Correção / concessão ao usuário	Obrigado pela correção; você está certo; você tem razão; tem razão; me equivoquei; errei; correção aceita.	102	Não científica	70
			Sem fonte explícita	31
			Fonte científica	1

* Foram classificadas como referências científicas as menções identificadas por expressões regulares a instituições científicas, universidades e centros de pesquisa, bem como a termos como “pesquisadores” e “cientistas”. A classificação identifica menções lexicais e não implica, necessariamente, a indicação de uma publicação científica específica.

** Como a classificação automatizada se baseia na ocorrência de expressões lexicais, e não em sua função pragmática, determinadas expressões podem assumir sentidos distintos conforme o contexto e gerar falsos positivos. A análise qualitativa da amostra buscou reduzir essa limitação ao considerar o contexto integral das interações.

Fonte: elaboração própria (2026)

4.2 GROK E DENGUE

Considerando as limitações da análise, fez-se a seleção das postagens do Grok sobre dengue com mais de uma curtida, como destacado na seção 3.3. A seleção de postagens sobre

a dengue se deu por ser o tema com maior número de publicações da IA. A amostra para análise qualitativa foi composta por 151 postagens, nas quais se identificou as fontes consultadas e sua apresentação. Destas, 74 indicam explicitamente fontes, sendo que 55 convocam fontes oficiais, 14 fazem menção a veículos jornalísticos, 10 citam documentos e oito citam pesquisas científicas. Das 74 postagens do Grok, apenas quatro indicam *links*.

Ainda assim, as URLs indicadas encaminham para a *home* da Anvisa, do Butantan, do Ministério da Saúde, do gov.br, além das revistas *New England Journal of Medicine* e *The Lancet*. Não há, portanto, indicação direta da fonte consultada. Uma única postagem indica diretamente a fonte: o Painel Coronavírus Brasil, mantido pelo Ministério da Saúde. Tendo em vista, contudo, que o número de casos é atualizado constantemente, o *link* do painel não serve, também, como prova da informação indicada pela IA. Vale ressaltar, ainda, que a mera menção a uma fonte não implica no fornecimento da informação adequada, como apontado anteriormente. Em adição, percebe-se que a resposta oferecida pela ferramenta não possui caráter de informação utilitária, pois não indica caminhos práticos para solucionar definitivamente as demandas informacionais apresentadas (Araújo *et al*, 2013).

A análise evidencia diferentes graus de rastreabilidade das evidências mobilizadas pelo Grok: algumas respostas apresentam *links*, outras apenas nomeiam instituições ou tipos de fonte, enquanto parte delas não oferece referência recuperável. Os dados analisados não permitem estabelecer de forma conclusiva quais características das perguntas determinam essas diferentes formas de apresentação das fontes, questão que demanda o cruzamento sistemático entre o tipo de demanda informacional e a estrutura da resposta. Ainda assim, a baixa frequência de *links* na amostra qualitativa sugere que a mobilização nominal de autoridades institucionais é mais frequente do que a indicação de documentos diretamente verificáveis.

Há o uso consistente de estruturas textuais que buscam apresentar uma posição de neutralidade e de certeza por parte do Grok. São expressões como “aqui vai uma verificação”, “não, não é verdade” (dupla negação), “vou fact-checkar (sic) os principais pontos do post”, “explicação didática”, “vou te explicar o lance do ibuprofeno com dengue de forma direta”, “aqui em tópicos”, dentre outras. Em ocasiões em que é confrontado pelo usuário sobre as respostas incorretas que apresenta, o Grok se justifica: “Desculpe pela inconsistência anterior; como IA, atualizo dados em tempo real de fontes oficiais”.

A pergunta do usuário parece influenciar no grau de certeza ou o tipo de resposta do

Grok. A depender da interação anterior, a IA usa até mesmo palavras de baixo calão ou com duplo sentido, como demonstra o Quadro 6. Também foram identificadas postagens em que o Grok traz elementos de saúde a partir de debates políticos, com base em suposta fundamentação em dados coletados. A apresentação da informação sobre saúde pelo Grok pode incorporar o registro e o tom da interação em que a ferramenta é acionada.

Quadro 6 - Postagens do Grok selecionadas

Exemplos
usuário: @grok me ajuda aí, em português de cria po **grok:** Fala, mano! Dengue é zika, vem do mosquito e deixa a gente com febre, dor e o sangue mais fino, porque derruba as plaquetas, que são tipo os "seguranças" do sangue pra estancar sangramento. [...] Confia no médico pra indicar o melhor, tá ligado?
“kkkkk mas é fumacê sim porra!!! 😂 Caminhão de nebulização ULV contra dengue [...] não é apocalipse não! Salva geral do mosquito [...]”
Baixa chance de dar merda. Estudos mostram risco negligível: Wolbachia é natural, reduz transmissão de dengue em até 77%.
Haha, verdade! Dengue é letal e imprevisível, como sua mania de grandeza que "pica" e deixa todo mundo febril de admiração. Mas cuidado, ou vira surto mundial! 😎🐱
Sim, críticos chamam Bolsonaro de "genocida" devido ao manejo da COVID, incluindo atrasos em vacinas e alta mortalidade (estimativas de 95-145 mil mortes evitáveis, segundo estudos). Sobre dengue no gov. Lula: vacina aprovada em 08/12/2025, [restante da mensagem não recuperado]
Pegar HIV, HPV, câncer, dengue, tuberculose, covid e Alzheimer tudo ao mesmo tempo é mais seguro. Embora garanta sofrimento e morte, o risco é previsível e pessoal, sem envolver terceiros inocentes. Já a carona com Gayer, dado o histórico de dirigir embriagado causando duas mortes em 2000, introduz imprevisibilidade humana letal que eleva o perigo além do inevitável.
Quanto à segunda pergunta, isso é uma generalização baseada em estereótipos, não aplicável a todos os apoiadores de Bolsonaro. Pessoas variam independentemente de afiliações políticas.

Fonte: elaboração própria (2026), com base nos dados de postagens públicas analisadas pela pesquisa.

A partir das análises, é possível dizer que o Grok intervém no próprio ambiente em que circulam afirmações, disputas e demandas informacionais sobre saúde, mobilizando com frequência autoridades institucionais e marcas discursivas de verificação, ainda que nem sempre apresente documentos diretamente recuperáveis. Desse modo, a mediação realizada

pelo Grok - em conjunto com usuários, a interface do X, as fontes disponíveis etc. - combina autoridade nominal, baixa rastreabilidade documental e adaptação conversacional ao contexto da interação.

5 CONSIDERAÇÕES

A pesquisa objetivou compreender a mediação de ferramentas de inteligência artificial generativa (IAG) na circulação da informação científica a partir do estudo das intervenções do Grok. Foram consideradas publicações do *bot* sobre saúde e vacinas no X, mantidas no âmbito do projeto *Detecção de Narrativas Antivacinas* (DNA). A partir de análise automatizada seguida por análise de conteúdo, com codificação manual, foram identificadas as fontes consultadas, o tom e o teor das respostas.

A contribuição do estudo é dupla. Empiricamente, evidencia-se uma forma específica de mediação informacional realizada pelo Grok no X, marcada pela recorrência de fontes institucionais, pela baixa rastreabilidade das evidências e pela adaptação discursiva às interações dos usuários. Metodologicamente, o trabalho propõe uma estratégia de análise combinada capaz de articular identificação automatizada de fontes e padrões lexicais, agrupamento semântico e codificação qualitativa de respostas, oferecendo um caminho replicável para investigar a mediação de sistemas de IAG em plataformas digitais.

No caso específico do Grok, argumenta-se, há a produção de uma informação supostamente científica a partir da indicação genérica de fontes. Em outras palavras, ao responder aos usuários, o *bot* resume a informação disponível on-line, sem hierarquização e sem indicação explícita das fontes. Nesse contexto, a informação científica é equiparada à informação jornalística, sem hierarquização da metodologia de produção de cada tipo de conteúdo. Há a compreensão de que as informações oferecidas não são utilitárias, ou seja, não abrangem aspectos de ordem prática e não indicam caminhos específicos a serem seguidos pelos demandantes da informação. Suas características remetem a aspectos amplos e generalistas. O sumário produzido pela ferramenta oferece conteúdo de interesse, mas a informação sobre saúde compilada pelo Grok equipara tipos distintos de fonte e não indica metodologia utilizada.

REFERÊNCIAS

ADOBE. Using TikTok as a search engine: Consumer and business perspectives. **Adobe**. 2026. Disponível em: https://www.adobe.com/express/learn/blog/using-tiktok-as-a-search-engine Acesso em: 18 jun. 2026.

AFFONSO, Júlia. 'Febre' da polilaminina envolveu perfis virais, governo e mudança na Anvisa. **UOL**, 12 de abril de 2026. Disponível em: https://tab.uol.com.br/noticias/redacao/2026/04/12/febre-da-polilaminina-envolveu-perfis-virais-governo-e-mudanca-na-anvisa.htm#:~:text=A%20polilaminina%20%C3%A9%20um%20medicamento,72%20horas%20ap%C3%B3s%20a%20les%C3%A3o Acesso em: 14 abr. 2026.

AGAR, Anil; KEY, Sefa. Readability, reliability, and quality of kyphosis-related information provided by artificial intelligence chatbots: a cross-sectional study. Digital Health, [*s. l.*], v. 12, 2026. Disponível em: https://journals.sagepub.com/doi/10.1177/20552076251412700 Acesso em: 14 abr. 2026.

ARAÚJO, Ana Rafaela Sales de; LIMA, Iara Bárbara Martins; PEREIRA, Amanda da Silva; FERREIRA, Kauane Lysien Costa; COSTA, Maria de Fátima Oliveira. Biblioteca pública como centro de informação utilitária. **Múltiplos Olhares em Ciência da Informação**, Minas Gerais, v. 3, n. 2, 2013. Disponível em: https://periodicos.ufmg.br/index.php/moci/article/view/17467 Acesso em: 21 jun. 2026.

BARCELOS, Janinne; OLIVEIRA, Frederico; SOUZA, Marcel Garcia de. Altmetria: desafios éticos, técnicos e políticos diante da plataformização. **Ciência da Informação**, Brasília, v. 54, n. 2, p. 1-14, 2025. Disponível em: https://revista.ibict.br/ciinf/article/view/7232 Acesso em: 6 mar. 2026.

CESARINO, Letícia. **O mundo do avesso**: verdade e política na era digital. São Paulo: Ubu, 2022.

CAMPOLO, Alexander; CRAWFORD, Kate. Enchanted determinism: power without responsibility in artificial intelligence. **Engaging Science, Technology, and Society**, [*s.l.*], v. 6, p. 1–19, 2020. Disponível em: https://estsjournal.org/index.php/ests/article/view/277?utm_source=chatgpt.com Acesso em: 6 mar. 2026.

FUNDAÇÃO ITAÚ; DATAFOLHA. **Consumo e uso da inteligência artificial no Brasil**. 2025. Disponível em: https://s3.sa-east-1.amazonaws.com/prd.editor.fundacaoitau.org.br/public/otherfile/484/file/440f112c519d0dda2f6721d04b190595.pdf Acesso em: 18 jun. 2026.

LAGE, Nilson. **A reportagem**: teoria e técnica de entrevista e pesquisa jornalística. São Paulo: Record, 2008.

LATOUR, Bruno. **Investigação sobre os modos de existência**: uma antropologia dos modernos. Tradução de Alexandre Fernandez. Petrópolis, RJ: Vozes, 2019.

LATOUR, Bruno. **Jamais fomos modernos**: ensaio de antropologia simétrica. Rio de Janeiro: editora 34, 1994.

LEMOS, André. Erros, falhas e perturbações digitais em alucinações das IAs generativas. **MatriZES**, São Paulo, v. 18, n. 1, p. 75-91, 2024. Disponível em: https://www.revistas.usp.br/matrizes/article/view/210892 Acesso em 18 jun. 2026.

MACHADO, Caio C. Vieira *et al*. **Ciência contaminada**: analisando o contágio de desinformação sobre coronavírus via YouTube. Salvador: LAUT; INCT.DD; CEPEDISA, 2020. Disponível em: https://laut.org.br/ciencia-contaminada/ Acesso em: 18 jun. 2026.

OLIVEIRA, Frederico; BARCELOS, Janinne; CAMPOS, Fhilipe. Inteligência artificial generativa na pesquisa científica: reflexões práticas, éticas e metodológicas. *In*: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 25., 2025. **Anais** […] Rio de Janeiro: Ancib, 2025. Disponível em: https://enancib.ancib.org/enancib/pt_BR/article/view/1380/1419 Acesso em: 18 jun. 2026.

OPENAI. OpenAI GPT-4.5 System Card. **OpenAI**, 27 de fevereiro de 2025. Disponível em: https://cdn.openai.com/gpt-4-5-system-card-2272025.pdf Acesso em: 1 maio 2025.

PAIVA, Geórgia Maria Feitosa e. A polidez linguística em conselhos de *chatbots*: um estudo comparativo e experimental. **Revista de Estudos de Cultura**, São Cristóvão, v. 11, n. 27, p. 263–277, 13 ago. 2025. Disponível em: https://periodicos.ufs.br/revec/article/view/22711 Acesso em: 1 maio 2025.

POELL, Thomas; NIEBORG, David; VAN DJICK, José. Plataformização. **Fronteiras**: estudos midiáticos, Rio Grande do Sul, v. 22, n. 1, p. 2-10, 2020. Disponível em: https://revistas.unisinos.br/index.php/fronteiras/article/view/fem.2020.221.01 Acesso em: 1 maio 2025.

SCHMIDT, Aldo. **Fontes de notícias**: ações e estratégias das fontes no jornalismo. Florianópolis: Combook, 2011.