

Verbetes Metodológico

Análise de conteúdo

Rafael Sampaio¹ 

¹Professor da Universidade Federal de Pernambuco, UFPE, Recife, PE, Brasil.

E-mail: cardososampaio@gmail.com

DOI: [10.17666/ebcs.2026.039](https://doi.org/10.17666/ebcs.2026.039)

PREÂMBULO

A análise de conteúdo é uma técnica de pesquisa que submete materiais escritos, verbais ou visuais a procedimentos sistemáticos e verificáveis por terceiros, a fim de produzir inferências sobre os contextos em que esses materiais foram produzidos, circularam e foram usados. Admite desenhos quantitativos, qualitativos ou mistos e foi desenvolvida para descrever padrões em grandes conjuntos de mensagens e investigar antecedentes e consequências prováveis da comunicação. Seu alcance tem limites claros: sozinha, ela não revela como públicos concretos receberam uma mensagem, quais eram as intenções de quem a produziu, nem o sentido integral de um texto, questões que exigem outros dados e métodos. No Brasil, a técnica é amplamente empregada nas humanidades e nas ciências da saúde, mas avaliações recentes apontam fragilidades nas etapas de treinamento, teste de confiabilidade e documentação, além de forte dependência de um único manual. O verbete trata da origem e dos fundamentos da técnica, de seus procedimentos, vantagens e limites, de aplicações brasileiras em ciência política, comunicação, relações internacionais e ensino de sociologia, e dos debates atuais sobre sua atualização.

SUMÁRIO

1. Histórico e fundamentos epistemológicos
2. Procedimentos principais
3. Vantagens, limites de validade e críticas conhecidas
4. Exemplos de aplicação em ciências sociais
5. Debates contemporâneos e variações
6. Referências bibliográficas

Aprovado em: 28/09/2026

Editor(a) responsável: Adriano Codato



Este é um artigo publicado em acesso aberto (Open Access) sob a licença Creative Commons Attribution (<https://creativecommons.org/licenses/by/4.0/>), que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Enciclopédia Brasileira de Ciências Sociais (EBCS). ANPOCS, 039 - 2026

1. Histórico e fundamentos epistemológicos

Práticas sistemáticas de leitura e contagem existem há séculos, sobretudo na classificação de textos religiosos segundo determinadas doutrinas, e estudos de jornais do fim do século XIX já contavam assuntos e mediam a extensão das matérias. Esses precedentes, no entanto, permaneciam dispersos e sem controle metodológico: a análise de conteúdo só se constituiu como técnica científica quando seleção, unidades, classificação e interpretação passaram a ser tratadas como partes de um mesmo desenho de pesquisa (Krippendorff, 1980/2018).

A análise de conteúdo ganhou forma científica na comunicação política diante da necessidade de comparar propaganda em grande escala. Harold Lasswell, precursor da versão moderna da técnica, examinou as técnicas de propaganda empregadas na Primeira Guerra Mundial em *“Propaganda Technique in the World War”*, de 1927, e, anos depois, coordenou pesquisas sobre comunicações de guerra na Biblioteca do Congresso dos Estados Unidos, nas quais mensagens de rádio eram classificadas por analistas treinados. Ali já estava o problema que permaneceria no centro da técnica: se pessoas diferentes atribuem códigos distintos ao mesmo material, a inferência depende de critérios capazes de tornar o desacordo observável e tratável. A origem militar e behaviorista explica a ênfase inicial na mensuração, mas não encerra a trajetória posterior.

Bernard Berelson condensou esse programa em *“Content Analysis in Communication Research”*, de 1952, ao definir a análise de conteúdo como descrição objetiva, sistemática e quantitativa do conteúdo manifesto da comunicação. A definição permitia comparar mensagens e registrar frequências, ao custo de restringir a análise ao que pudesse ser reconhecido na superfície textual. Quando a técnica passou a circular na psiquiatria, na antropologia e nos estudos de linguagem, a restrição ficou difícil de sustentar, porque contar ocorrências dizia pouco sobre ironia, ambiguidade, ausência ou relação entre texto e situação social (Bardin, 1977/2016).

O contraste entre conteúdo manifesto e latente nunca coincidiu com a separação entre quantidade e interpretação, porque classificar uma menção como favorável ou desfavorável já depende de regras sobre negação, ironia, citação e contexto. Conteúdos latentes podem ser codificados quando o conceito está bem definido, os indícios são observáveis no material e as regras reduzem a margem de decisão de quem codifica. O problema começa quando o analista toma a própria leitura como o sentido que estaria no texto. A objetividade deixou de ser uma leitura supostamente neutra e passou a ser a obrigação de explicar em público as escolhas feitas.

Krippendorff (1980/2018) sistematizou essa mudança ao definir a técnica pela produção de inferências válidas e replicáveis dos textos para os contextos de seu uso. O sentido deixa de ser propriedade depositada no material à espera de descoberta e passa a resultar de uma construção do pesquisador, que liga sinais observáveis, conceitos e contexto, razão pela qual nenhuma frequência fala por si. A validade indaga se a pesquisa sustenta a conclusão pretendida; a confiabilidade examina a estabilidade das classificações sob condições comparáveis; a replicabilidade permite que terceiros reconstruam o percurso e testem as decisões tomadas. A qualidade científica depende da relação entre esses três critérios.

Na tradição anglo-saxã, a técnica passou a seguir os princípios das ciências experimentais, com a exigência de que a codificação não dependa de quem codifica e de que os resultados sejam replicáveis. Daí vêm a amostragem probabilística, o livro de códigos, o treinamento, o teste de confiabilidade e os padrões de transparência que organizam suas fases (Lacy et al., 2015; Neuendorf, 2017).

A tradição qualitativa de língua alemã, sistematizada por Philipp Mayring e por Margrit Schreier, manteve a exigência de regras explícitas e as principais etapas da versão quantitativa, sem tomar a quantificação como destino obrigatório. Não existe, porém, uma única análise de conteúdo qualitativa. Schreier et al. (2019) mostram que o método se desenvolveu como híbrido no debate alemão e como termo guarda-chuva na literatura anglófona, com divergências sobre a exclusividade das categorias, o papel dos temas e os critérios de avaliação, já que uma tradição fala em confiabilidade e a outra em credibilidade.

Na França, Laurence Bardin publicou em 1977 sua proposição da técnica, marcada pela disputa com a semiologia e por uma perspectiva híbrida que aceita categorias indutivas e dedutivas, hoje mais próxima da análise temática. Sua obra se tornou a base da aplicação brasileira, relida por autoras como Maria Laura Franco, na educação, e Maria Cecília Minayo, na saúde coletiva, e foi frequentemente confundida com qualquer forma de análise qualitativa de dados. Essa adoção quase inquestionada encobriu a genealogia ligada a Lasswell e Berelson, reduziu o contato com o debate internacional e produziu uma monocultura de citação que manteve a técnica praticamente parada no Brasil, com pouca atenção a treinamento, livro de códigos, teste de confiabilidade e princípios de transparência (Sampaio et al., 2025).

2. Procedimentos principais

Na análise de conteúdo, o analista produz a medida, ou seja, define o que vai ser observado no material. Sua interpretação não pode ser retirada do processo, pois identificar uma palavra, uma imagem ou uma cena depende da pergunta e do conceito que orienta a classificação, razão pela qual a técnica é prioritariamente dedutiva e sempre semântica. O controle metodológico começa quando essas escolhas deixam de operar de forma tácita e passam a ser registradas em regras examináveis. O procedimento torna-se, assim, uma sequência de decisões encadeadas, em que cada escolha define o que o estudo poderá concluir e o que fica fora de seu alcance.

A constituição do *corpus* já contém uma hipótese sobre o fenômeno. Uma amostra probabilística sustenta inferências sobre uma população delimitada quando o universo de mensagens é conhecido e cada item tem probabilidade calculável de seleção, enquanto um *corpus* montado por escolha do pesquisador permite examinar casos relevantes, variações internas ou ocorrências raras, sem autorizar generalização estatística (Krippendorff, 1980/2018; Neuendorf, 2017). A semana construída, que sorteia dias de semanas diferentes para formar uma semana artificial, procura controlar ciclos editoriais, enquanto buscas digitais dependem de termos, interfaces e regras de acesso que também selecionam o material. Volume elevado não corrige população mal delimitada, e um censo do conteúdo acessível por uma plataforma continua restrito ao que ela tornou disponível.

As unidades tornam visível essa seleção. Uma edição de jornal, um programa eleitoral ou uma postagem pode constituir o item coletado, chamado unidade amostral. A codificação recai sobre a unidade de análise, que pode ser uma fala, um parágrafo ou uma cena, e consulta uma unidade de contexto composta pelo material necessário à interpretação daquele segmento. Uma mesma postagem pode ser unidade amostral e unidade de análise, enquanto a conversa anterior funciona como contexto. Confundir esses níveis altera denominadores, frequências e comparações, assim como impede a replicação (Sampaio & Lycarião, 2021).

Categorias e códigos traduzem a pergunta de pesquisa em distinções observáveis. Categorias dedutivas partem da teoria, de conceitos ou de questões práticas delimitadas, enquanto categorias indutivas surgem dos dados e podem ser revistas após contato sistemático com o material, embora a técnica funcione melhor com as primeiras. Exclusividade, exaustividade e homogeneidade são critérios úteis quando o desenho exige um só código por unidade e cobertura integral do material, que é o padrão mais adotado (Bardin, 1977/2016; Krippendorff, 1980/2018). O que precisa permanecer inequívoco é a relação entre cada código, o conceito que ele operacionaliza e a inferência que dele será derivada.

Por essa lógica de categorização semântica e pelo valor atribuído à replicação, todas as categorias e seus códigos precisam ser testados antes do estudo final. Um código de incivildade, por exemplo, precisa distinguir crítica, insulto, ironia e ameaça sem reunir toda linguagem negativa. Rodadas piloto localizam sobreposições, lacunas e regras que funcionam apenas para os exemplos conhecidos pela equipe, e nessa fase categorias podem ser fundidas ou abandonadas. A flexibilidade ajuda a construir o instrumento, que é fechado quando a codificação começa de fato (Neuendorf, 2017; Schreier et al., 2019).

O livro de códigos registra essa relação por meio de definições, critérios de inclusão e exclusão, exemplos e casos limítrofes. É etapa obrigatória para garantir codificação comparável dentro da equipe e para permitir que outras pessoas repitam o estudo, e só se considera concluído após resultado positivo no teste de confiabilidade. O treinamento usa uma amostra para localizar ambiguidades antes da codificação definitiva, enquanto o teste estima se a versão estabilizada gera classificações comparáveis sem negociação imediata. Mudanças posteriores pedem novo treinamento e novo teste, pois acordos obtidos com uma versão não validam a seguinte (Sampaio & Lycarião, 2021).

O teste de confiabilidade pede que todos os envolvidos codifiquem o mesmo material de forma independente, com posterior comparação por índices como o alfa de Krippendorff, aplicável a diferentes números de codificadores e níveis de mensuração, além do kappa de Cohen e do pi de Scott, usados sob condições mais restritas (Krippendorff, 1980/2018). Lacy et al. (2015) acrescentam recomendações que a produção brasileira raramente segue: ao menos um codificador não deve ter participado da construção do protocolo, porque a confiabilidade reside no instrumento e não em quem o escreveu, e o tamanho da amostra do teste deve resultar de cálculo, não de percentual do *corpus*.

Nas análises qualitativas, testes de confiabilidade continuam indicados, com diferenças relevantes. Parte da literatura segue o padrão quantitativo, enquanto outra defende uma medida descritiva e baseada em processo, que caracterize o grau de concordância entre codificadores independentes sem convertê-lo em coeficiente. São aceitas ainda formas de validação próprias da pesquisa qualitativa, como credibilidade, dependabilidade,

transferibilidade, saturação e triangulação entre pesquisadores (Schreier et al., 2019). O teste não é descartado, mas passa a integrar um conjunto maior de estratégias, no qual pesam mais o relato transparente dos procedimentos, a discussão documentada das divergências e o exame que a equipe faz das próprias escolhas.

A codificação final produz dados que podem ser tratados por frequências, associações, comparação entre grupos ou leitura interpretativa de ocorrências e ausências. Na pesquisa quantitativa, isso costuma levar ao teste de hipóteses ou ao cruzamento de variáveis, com qui-quadrado e resíduo padronizado para dados categoriais e análise de variância quando há variáveis de outro tipo (Neuendorf, 2017). Na vertente qualitativa, as frequências obtidas nos códigos são o ponto de partida, e o trabalho consiste em explicar os achados com excertos originais e reconstruir a história que as categorias contam (Schreier et al., 2019).

A replicação está no desenho da técnica desde sua formulação, de modo que os materiais da pesquisa devem ser tornados públicos, em especial o livro de códigos, o banco de dados codificado e, quando for eticamente possível, o *corpus* original. Só com esses elementos alguém consegue examinar o vínculo entre conceito, unidade, contexto e conclusão.

3. Vantagens, limites de validade e críticas conhecidas

Quando trabalha com material produzido independentemente da pesquisa, como jornais, programas eleitorais, atas, ofícios, sentenças judiciais, relatórios de organismos internacionais e postagens em redes sociais, a análise de conteúdo elimina o efeito de resposta que afeta entrevistas, grupos focais e questionários, porque ninguém escreveu aquele documento sabendo que ele seria codificado. A técnica também alcança atores que não se deixam entrevistar, como cúpulas partidárias, ministérios, tribunais e chefes de Estado, além de períodos que já não admitem coleta direta.

Como técnica híbrida, entre o qualitativo e o quantitativo, ela lida melhor com volume do que as demais análises de material textual baseadas em leitura humana. Depois do teste de confiabilidade, o *corpus* pode ser dividido entre os codificadores, e livros de códigos curtos, que podem ser aplicados com rapidez, ampliam ainda mais a escala (Krippendorff, 1980/2018). Análise de discurso e análise temática tentam esgotar os sentidos de cada conteúdo e gastam muito mais tempo por unidade. O custo financeiro também é baixo, já que boa parte dos *corpora* é pública e dispensa trabalho de campo, deslocamento e financiamento externo, o que faz diferença em programas de pós-graduação brasileiros com orçamento apertado.

Codificar milhares de unidades e poder voltar ao trecho de origem permite acompanhar padrões de representação e suas mudanças no tempo. Um mesmo livro de códigos aplicado a décadas de jornais mostra quais grupos sociais aparecem na cobertura, quem fala, que temas ganham espaço e quando um enquadramento antes dominante começa a desaparecer. Séries longas ligam essas mudanças a eventos datados, como eleições, decisões judiciais ou crises econômicas, o que um *corpus* pequeno não permite. Se a amostra for probabilística e a população de mensagens estiver bem delimitada, as conclusões valem para o conjunto das mensagens daquele período, não apenas para os textos que a equipe leu (Neuendorf, 2017).

Os dados codificados ainda se articulam com facilidade a outros bancos. Uma variável de ataque em propaganda eleitoral pode ser cruzada com séries de intenção de voto, uma escala ideológica construída a partir de programas partidários pode ser comparada a votações nominais no Legislativo e a frequência de um tema em discursos diplomáticos pode ser lida contra fluxos de comércio ou votos em organismos multilaterais. Poucas técnicas de análise textual produzem material tão combinável com *surveys*, dados administrativos e registros eleitorais, e é normalmente essa combinação que sustenta a inferência causal que a análise de conteúdo sozinha não alcança.

Entre as análises qualitativas, ela é a mais replicável e a mais falseável. O livro de códigos, as regras de aplicação, a unidade de análise e os critérios de amostragem ficam escritos antes da codificação, de modo que outra equipe pode tomar o mesmo material, aplicar o mesmo instrumento e comparar seus resultados com os originais. Lacy et al. (2015) formulam o ponto com precisão ao lembrar que a confiabilidade reside no protocolo e não nos codificadores, o que obriga o protocolo a funcionar longe de quem o criou. O teste de confiabilidade torna esse controle visível, porque mostra em que categorias houve acordo e em quais a regra falhou (Sampaio & Lycarião, 2021). Em análise de discurso ou análise temática, a leitura depende de quem analisa e dificilmente é repetida nos mesmos termos.

Essa disciplina tem um efeito cumulativo pouco mencionado, já que instrumentos publicados podem ser reaplicados a outro país, outro período ou outra plataforma, e é assim que se constroem séries comparadas de longo prazo. Lacy et al. (2015) chegam a recomendar que se relatem as variáveis que falharam no teste e foram descartadas, porque saber o que não funcionou poupa trabalho alheio. Construir um livro de códigos obriga ainda o pesquisador a transformar um conceito abstrato em algo observável, exigência que costuma melhorar o desenho mesmo quando o estudo final é interpretativo.

A principal desvantagem da técnica está na decisão de valorizar objetividade e replicabilidade sobre a validade do que se mede. Transformar textos, imagens ou cenas em códigos custa as relações internas do material, que costumam ser justamente o que produz sentido, e a contagem diz quantas vezes algo foi classificado em cada categoria sem dizer se aquilo teve importância social, qual foi a intenção por trás e que efeito gerou em quem leu ou assistiu. Frequência também não mede peso, porque um termo repetido em todas as atas pode ser fórmula protocolar, enquanto uma menção única pode mudar o rumo de uma decisão.

Categorias endurecem distinções que são do pesquisador e não do material. Classificar partidos em esquerda e direita, ou falas em civis e incivis, cria fronteiras que o material raramente sustenta com nitidez. Códigos residuais escondem os casos que desafiam a teoria, enquanto categorias amplas fabricam semelhança onde havia diferença. Na comparação entre países, períodos ou grupos sociais, uma regra estável pode significar coisas distintas, porque equivalência formal não garante equivalência cultural.

O tamanho da unidade acrescenta outro dilema, já que segmentos pequenos perdem contexto e segmentos grandes dificultam códigos precisos. O que não foi dito escapa a qualquer desenho que olhe somente para o que aparece, e o mesmo vale para o que nem chega ao *corpus*, porque arquivos históricos têm lacunas, plataformas mudam regras de acesso e buscas por palavras-chave selecionam material sem que o pesquisador perceba. Lacy et al. (2015) relatam que 91% dos resultados de uma busca por nomes de países no *New York Times*

não tinham aquele país como assunto principal e recomendam medir a precisão e a cobertura da expressão de busca antes de tratar o resultado como *corpus*.

O pesquisador decide o tempo todo entre categorias complexas, que distinguem melhor o fenômeno e são mais difíceis de replicar, e categorias simples, que geram mais acordo ao custo de apagar diferenças que importam. Um índice alto de concordância mostra apenas que a equipe classificou de modo parecido; o problema começa quando esse número encerra a discussão sem que ninguém pergunte se as categorias mediam o que interessava e quais casos ficaram fora do instrumento (Sampaio & Lycarião, 2021). Treinamento muito dirigido ensina os codificadores a repetir a mesma leitura fraca e a hierarquia da equipe transforma a opinião da pessoa mais experiente em consenso aparente (Krippendorff, 1980/2018).

Relatar bem não conserta a pesquisa, já que um estudo pode errar e descrever o erro com precisão. Omitir unidades, livro de códigos, treinamento, divergências e resultados por categoria, porém, impede que o leitor encontre esses erros. No Brasil, a distância entre usar e documentar é grande. Entre 549 artigos que aplicaram a análise de conteúdo qualitativa, 2,6% informaram teste de confiabilidade, 0,5% mencionaram treinamento de codificadores, 7,3% relataram alguma forma de validação e 2,4% adotaram alguma forma de transparência metodológica; Bardin apareceu em 66,8% dos textos e nenhuma das 367 citações examinadas fez contraste crítico com a obra (Sampaio et al., 2022). Entre 71 tutoriais brasileiros, todos citaram Bardin, 62% repetiram o modelo de três fases e a média ficou em 2,85 de nove indicadores de qualidade (Sampaio et al., 2025). Os números associam concentração bibliográfica e pouca diversidade de critérios, sem atribuir todas as falhas a uma só obra. Como as evidências vêm do Brasil, estender o diagnóstico ao restante da América Latina exige investigação comparável.

4. Exemplos de aplicação em ciências sociais

Nas ciências sociais, a análise de conteúdo pode ser aplicada a uma grande variedade de objetos, porque permite examinar materiais escritos, visuais, sonoros e audiovisuais produzidos em diferentes contextos sociais e institucionais. Seu uso abrange desde artigos científicos, jornais, editoriais, leis, documentos oficiais e comunicações institucionais até entrevistas, grupos focais, registros etnográficos e materiais didáticos. Na prática, a técnica pode ser empregada na análise de programas partidários, planos de governo, pronunciamentos, debates eleitorais e propaganda política, inclusive a do horário eleitoral gratuito. No ambiente digital, pode orientar pesquisas sobre postagens em redes sociais, memes, comentários em portais de notícias, conversações públicas, processos deliberativos e manifestações de desrespeito. Essa diversidade permite investigar temas, enquadramentos, argumentos, valores, representações, estratégias discursivas e padrões de interação em objetos muito distintos.

Borba (2019) perguntou se a propaganda negativa na eleição presidencial de 2014 variava entre canais de comunicação. O *corpus* reuniu 4.098 mensagens dos três principais candidatos, distribuídas entre *spots* de televisão e de rádio, falas em debates, declarações publicadas em “O Globo” e exibidas no “Jornal Nacional” e postagens no *Facebook*, codificadas

nas categorias de aclamação, ataque e defesa. Os ataques predominaram nos debates, em que responderam por cerca de 55% das mensagens, foram frequentes nas declarações publicadas no jornal e nos *spots* de televisão e raros no Facebook e no telejornal. A negatividade varia, portanto, conforme o controle do candidato sobre a mensagem, a reciprocidade do confronto e a possibilidade de identificar quem ataca.

Novelli e Pereira (2019) investigaram se os Estados banhados pelo Atlântico Sul tratam aquele espaço oceânico como uma região. A busca pela expressão “*South Atlantic*” na biblioteca digital da Organização das Nações Unidas (ONU) retornou 177 discursos proferidos na Assembleia Geral entre 1986 e 2015 por 43 Estados. O *corpus* final ficou com os 106 discursos dos 24 membros da Zona de Paz e Cooperação do Atlântico Sul (ZOPACAS). Os dados indicam que o projeto regional teve força entre 1986 e 1994, perdeu fôlego até 1998 e não se recuperou, de modo que o Atlântico Sul segue lido como espaço de interação entre a América do Sul e a África Ocidental.

Oliveira (2025) investigou como estudantes de uma escola estadual de ensino médio integral da periferia de Maceió compreenderam conceitos de Pierre Bourdieu após três sequências didáticas baseadas no filme “Que horas ela volta?”. As atividades abordaram disposições e capitais na primeira série, campo e espaço social na segunda e trajetórias e condições de possibilidade na terceira. Depois das aulas, dezesseis estudantes responderam voluntária e anonimamente a um questionário aberto on-line, cinco da primeira série, nove da segunda e dois da terceira. As respostas foram submetidas à categorização temática da análise de conteúdo e organizadas em eixos como posição no espaço social, modos de pensar, condições de agência, reprodução social, resistência reflexiva e desigualdade de classe. Os estudantes mobilizaram conceitos como capital, *habitus*, campo, trajetória e posição social para interpretar suas próprias vidas e relações familiares. Parte das respostas expressou otimismo sobre a possibilidade de romper ciclos sociais, com dezessete trechos classificados como orientados à transformação, enquanto treze manifestaram pessimismo e angústia diante da reprodução das desigualdades. O estudo mostra como a análise de conteúdo pode ser usada para identificar formas distintas de apropriação de uma teoria sociológica e avaliar os efeitos reflexivos de uma intervenção pedagógica.

5. Debates contemporâneos e variações

O debate mais urgente sobre a análise de conteúdo no Brasil não trata de inovação, mas de atraso. A técnica que aqui se pratica está ancorada em um manual publicado em 1977 e sem atualização substancial há mais de duas décadas, enquanto a discussão internacional se moveu bastante no mesmo período (Sampaio et al., 2025). Lacy et al. (2015) consolidaram quatorze recomendações de boas práticas que cobrem amostragem em bases digitais, número de codificadores, tamanho da amostra do teste, escolha de coeficientes e relato das variáveis descartadas.

A consequência prática é uma agenda de confiabilidade que já deveria estar incorporada às pesquisas brasileiras. Às duas recomendações mencionadas na seção 2, sobre o codificador externo à construção do protocolo e o cálculo do tamanho da amostra do teste, soma-se a de

relatar mais de um coeficiente, porque categorias raras produzem alfa baixo mesmo quando a concordância simples é alta (Lacy et al., 2015). Nenhuma dessas exigências depende de infraestrutura computacional, financiamento adicional ou treinamento estatístico avançado, e é por elas que a modernização da técnica no Brasil deveria começar.

Na vertente qualitativa, Schreier et al. (2019) mostram que não existe uma única análise de conteúdo qualitativa, mas várias, e que a divergência entre o debate alemão e o internacional alcança a própria definição de qualidade, com uma tradição falando em confiabilidade e a outra em credibilidade (*trustworthiness*). A documentação insuficiente dos procedimentos aparece nos dois lados como problema comum. Os autores propõem uma saída que não exige abandonar o rigor nem importar o aparato estatístico, porque a confiabilidade entre codificadores pode ser descritiva e baseada em processo. A recomendação envolve pelo menos dois codificadores, um deles afastado da coleta, um mesmo quadro analítico, discussão documentada das divergências e um livro de códigos construído por consenso e depois aplicado ao restante do material. O que se pede é registro, não coeficiente.

Programas de análise qualitativa assistida, como NVivo, Atlas.ti e MAXQDA, ajudam essa agenda quando são usados como registro, porque preservam o vínculo entre código e trecho, guardam as anotações de quem codificou e mostram quando uma regra mudou, quem a alterou e que parte do *corpus* foi afetada. Organizar documentos, aceitar ocorrências sugeridas pelo programa e delegar-lhe classificações, operações que convivem na mesma interface, transferem ao programa graus diferentes de julgamento.

A conexão com a análise textual automatizada (*text as data*) é o debate mais recente e convém tratá-la como complemento, já que a análise automatizada tende a ser basicamente léxica e computacional, enquanto a análise de conteúdo é semântica e se baseia na codificação humana. Lacy et al. (2015) já defendiam que a codificação algorítmica constitui método distinto, com erros próprios, e que validá-la apenas contra codificadores humanos supõe que os dois procedimentos produzam o mesmo dado.

Recentemente, modelos de inteligência artificial generativa levantaram novos debates sobre classificações automáticas e híbridas (humano-máquina). Ziems et al. (2024) avaliaram treze modelos de linguagem em 25 tarefas representativas de ciência social computacional e encontraram, nas tarefas de classificação, desempenho inferior ao de classificadores treinados e concordância apenas razoável com anotadores humanos; e, nas tarefas de geração, explicações que muitas vezes superam as produzidas por trabalhadores contratados para anotar. A recomendação dos autores é usar o modelo dentro da equipe de anotação, e não no lugar dela. Com o avanço dos modelos de linguagem, é provável que esse cenário mude.

Há também um uso menos óbvio, que trata a hesitação da máquina como dado. Costa e Bem-Haja (2025) propõem que a incerteza do modelo seja registrada em vez de silenciada. Quando o sistema atribui a um trecho pontuações próximas em duas categorias, como os 72% de empatia e os 68% de permissividade que encontraram em redações escolares, o desacordo aponta para uma fronteira conceitual mal resolvida no livro de códigos, e não para uma falha técnica a corrigir. O pesquisador escreve então um comentário curto explicando por que discorda, e esse comentário passa a integrar a trilha analítica. Os autores relatam ainda o que o sistema não viu, como o peso simbólico da exclusão espacial, o desvio em relação a expectativas de gênero e, sobretudo, a ausência de um tema como informação.

Quem usa esses modelos precisa dizer qual usou, registrando o nome e a versão do modelo, a data da execução, a instrução (*prompt*) escrita, os parâmetros escolhidos, quantas vezes a operação foi repetida e o que a equipe fez quando as respostas divergiram entre si. Como esses modelos aprenderam sobretudo com texto em inglês, ironia, xingamento e referência local podem ser mal classificados em um *corpus* brasileiro, além de que um material sensível enviado a um serviço externo levanta questões de guarda dos dados, jurisdição e identificação de quem falou.

6. Referências bibliográficas

Textos fundadores e genealogia

- Bardin, Laurence. (2016). *Análise de conteúdo* (Luís Antero Reto & Augusto Pinheiro, Trans.). Edições 70. (Obra original publicada em 1977).
- Krippendorff, Klaus. (2018). *Content analysis: An introduction to its methodology* (4th ed.). SAGE. (Obra original publicada em 1980). <https://doi.org/10.4135/9781071878781>.

Manuais e guias de referência

- Neuendorf, Kimberly A. (2017). *The content analysis guidebook* (2nd ed.). SAGE. <https://doi.org/10.4135/9781071873045>.
- Sampaio, Rafael Cardoso, & Lycarião, Diógenes. (2021). *Análise de conteúdo categorial: Manual de aplicação*. Enap. <https://repositorio.enap.gov.br/handle/1/6542>.

Crítica e debates metodológicos

- Costa, Antônio Pedro, & Bem-Haja, Pedro. (2025). Reflexive uncertainty AI for qualitative data analysis. *CEUR Workshop Proceedings*, 4114, 13–23. https://ceur-ws.org/Vol-4114/7_paper.pdf.
- Lacy, Stephen, Watson, Brendan R., Riffe, Daniel, & Lovejoy, Jennette. (2015). Issues and best practices in content analysis. *Journalism & Mass Communication Quarterly*, 92(4), 791–811. <https://doi.org/10.1177/1077699015607338>.
- Sampaio, Rafael Cardoso, Sanchez, Cristiane Sinimbu, Marioto, Djiovanni Jonas França, Araujo, Beatrice Cristina dos Santos, Herédia, Larissa Helena Olivares, Paz, Felipe Schwarzer, Tigrinho, Camila Schiavon, & Souza, Josiane Ribeiro de. (2022). Muita Bardin, pouca qualidade: Uma avaliação sobre as análises de conteúdo qualitativas no Brasil. *Revista Pesquisa Qualitativa*, 10(25), 464–494. <https://doi.org/10.33361/RPQ.2022.v.10.n.25.547>.
- Sampaio, Rafael Cardoso, Sanchez, Cristiane Sinimbu, Tigrinho, Camila Schiavon, Ernsen, Karina, Souza, Josiane Ribeiro de, & Duarte, Elysangela Dittz. (2025). *Chega de Bardin! O círculo vicioso da análise de conteúdo brasileira* [Preprint]. SciELO Preprints. <https://doi.org/10.1590/SciELOPreprints.12029>.
- Schreier, Margrit, Stamann, Christoph, Janssen, Markus, Dahl, Thomas, & Whittal, Amanda. (2019). Qualitative content analysis: Conceptualizations and challenges in research practice — Introduction to the FQS special issue “Qualitative Content Analysis I”. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 20(3), Artigo 38. <https://doi.org/10.17169/fqs-20.3.3393>.
- Ziems, Caleb, Held, William, Shaikh, Omar, Chen, Jiaao, Zhang, Zhehao, & Yang, Diyi. (2024). Can large language models transform computational social science? *Computational Linguistics*, 50(1), 237–291. https://doi.org/10.1162/coli_a_00502.

Aplicações empíricas

- Borba, Felipe. (2019). Medindo a propaganda negativa na TV, rádio, debates, imprensa e Facebook: O caso das eleições presidenciais de 2014. *Intercom: Revista Brasileira de Ciências da Comunicação*, 42(1), 37–56. <https://doi.org/10.1590/1809-5844201912>.
- Novelli, Douglas Henrique, & Pereira, Alexsandro Eugenio. (2019). O Atlântico Sul nos discursos na ONU: Uma análise de conteúdo sobre o caráter regional do espaço oceânico (1986–2015). *Conjuntura Global*, 8(2), 167–195. <https://doi.org/10.5380/cg.v8i2.70127>.

Oliveira, Francisco Érick de. (2025). Pierre Bourdieu e o cinema nacional no ensino médio: Percepções discentes e ensino de sociologia. *Revista TOMO*, 44, e22590. <https://doi.org/10.21669/tomo.v44.22590>.

VER TAMBÉM:

[Análise de discursos](#) | [Análise de texto automatizada/ *Text as data*](#) | [Confiabilidade](#) | [Validade](#) | [Pesquisa em arquivos](#) | [Etnografia digital](#) | [Métodos Mistos](#)

COMO CITAR ESTE VERBETE:

Sampaio, Rafael. (2026). Análise de conteúdo. In *Enciclopédia Brasileira de Ciências Sociais (EBCS)*. ANPOCS. 10.17666/ebcs.2026.039.