em questão
E-ISSN 1808-5245

A desinformação como um problema sistêmico: contribuições para uma reflexão epistemológica sobre o fenômeno

Leonardo Ripoll [I]
https://orcid.org/0000-0002-1345-1631

[I] Universidade de São Paulo, São Paulo, SP, Brasil

Resumo: O texto busca contribuir para a compreensão da desinformação contemporânea enquanto um fenômeno sistêmico, conforme apontam pesquisas recentes no campo dos estudos sobre desinformação (*disinformation studies*). Caracterizado pelo ambiente das plataformas digitais, argumentamos que esse fenômeno é intensificado pelo desenvolvimento das tecnologias cognitivas, atualmente representadas pela inteligência artificial generativa e pela emergente inteligência artificial agêntica. Para sustentar essa análise, relacionamos a origem terminológica da desinformação ao contexto da virada informacional no pensamento científico e filosófico, além de discutir concepções contemporâneas do tema presentes tanto na literatura científica quanto em documentos de organizações internacionais. Argumentamos que a desinformação transcende sua definição etimológica e se consolida como instrumento político e econômico, explorado em dinâmicas estruturadas pelas *big techs* no domínio do ciberespaço. Por fim, destacamos a urgência de reavaliar as relações informacionais e comunicacionais diante do risco de uma "pós-realidade" e defendemos a relevância da perspectiva semiótica para compreender os processos de produção de sentido que atravessam os fluxos de informação e desinformação.

Palavras-chave: desinformação; plataformas digitais; inteligência artificial; *fake news*; filosofia da informação

1 Introdução

O fenômeno da desinformação contemporânea atravessa todos os campos do conhecimento e afeta o cotidiano da esfera pública em seus âmbitos mais elementares: social, político, ambiental, sanitário, educacional, cultural e tecnológico. Marcado atualmente pelas dinâmicas provenientes das interações sociais em rede e impulsionado pelas plataformas digitais, o fenômeno foi terminologicamente constituído no início do século passado, a partir das estratégias infocomunicacionais evidenciadas nas guerras e regimes de governo que integram um cenário geopolítico de guerra informacional (*information warfare*), segundo autores como Acampa (2024) e Salaverría e Cardoso (2023).

É seguro considerar que a desinformação, embora antiga enquanto ideia, torna-se efetivamente um objeto de ampla discussão somente a partir do contexto digital das redes. A disseminação do termo, após a consolidação das mídias sociais, mobilizou um significado direcionado aos artefatos informacionais e ao fluxo em rede que caracteriza o ambiente digital contemporâneo. No início deste cenário, podemos admitir que havia alguma convergência em sinalizar a desinformação enquanto sinônimo de um objeto singular, passível de identificação e correção. Isso porque a primeira forma de desinformação que ficou largamente conhecida neste novo contexto foram as *fake news* – conteúdos falsos e anônimos que mimetizam o formato jornalístico e que surgiram a partir da facilidade de criação de *blogs* possibilitada pela chamada "web 2.0". Nesse momento, a desinformação era socialmente percebida como uma anomalia, algo ainda em caráter de exceção, mesmo que presente com crescente constância em um cotidiano que lidava com a transição das informações impressas para o texto em tela.

Mas os tipos de conteúdos desinformativos logo se diversificaram e se expandiram, conforme o cenário tecnológico e econômico permitiu um processamento de dados em maior quantidade e com maior velocidade. Pouco depois, as consequências da modulação algorítmica dentro da economia da atenção instaurada pelas plataformas digitais fizeram a desinformação se tornar um modelo de negócio e ser incorporada na retórica política digital. Essa escalada tecnológica permitiu sistemas de "dataficação" da vida – entendido aqui, a partir

de van Dijck (2014), como o processo que, com a chegada do *big data*, possibilitou o perfilamento humano para exploração econômica e influência política da opinião pública, estabelecendo um *modus operandi* em que disponibilizar informações da vida privada às empresas, de forma geral, se torna um novo modo de funcionamento social. Desde então, o tecido democrático sofre constantes rupturas e as ações de desinformação se tornaram estruturadas e mobilizadas pelas próprias empresas que detêm não somente o domínio dos dados pessoais, mas também as regras do ambiente digital em que a sociedade habita.

Nesse sentido, a desinformação contemporânea passou a ser diagnosticada como um fenômeno intrinsecamente conectado às plataformas digitais, as quais fornecem vantagens estruturais para a sua manutenção e amplificação (Recuero, 2025; Diaz Ruiz, 2025; Schneider, 2022), sem desconsiderar, entretanto, fatores externos ao digital que também contribuem para a sua sustentação. Torna-se, portanto, um sistema ativo e dinâmico, fomentado por incentivos econômicos, interesses políticos e por uma infraestrutura digital que se favorece com a sua permanência.

Com a chegada das tecnologias cognitivas (sobretudo o que começou a ser chamado de inteligência artificial generativa), a desinformação atingiu níveis refinados de elaboração, cujos efeitos sociais e políticos se mostram cada vez mais profundos. Estudos recentes têm ressaltado as questões sistêmicas e estruturais do fenômeno, que não pode mais ser compreendido fora de uma rede de ações e atores sociais responsável pela criação de um ecossistema que, com certa frequência, é alimentado por campanhas de desinformação provenientes de disputas geopolíticas (Diaz Ruiz, 2025; Recuero, 2024, 2025; Murphy *et al*., 2023; Salaverría; Cardoso, 2023; Figueiredo; Ferrari; Lima, 2023; Schneider, 2022; Dourado, 2021). De certa forma, o atual estado do fenômeno parece explicitar a sua origem histórica enquanto estratégia de poder diante de um cenário renovado de *information warfare* que se apresenta globalmente (Schroeder *et al*., 2026).

O objetivo deste artigo é trazer reflexões por meio de um breve panorama sobre esse processo, considerando o atual estágio do cenário informacional contemporâneo e os apontamentos encontrados nos estudos sobre desinformação

(*disinformation studies*), campo epistemológico em constituição. Trata-se, assim, de uma investigação teórica exploratória, parte de tese doutoral em andamento, em que os procedimentos metodológicos empregados foram direcionados a uma pesquisa bibliográfica sobre o tema da desinformação, com foco nas áreas da Comunicação e Informação.

A delimitação dos eventos que constituem a desinformação inclui tópicos já muito debatidos, como *fake news*, pós-verdade e *deepfakes*, mas também abrange um ecossistema que contempla diversos elementos envolvidos na produção, disseminação e recepção de mensagens e conteúdos. Esse ecossistema desenvolveu-se em consonância com os avanços científicos e tecnológicos iniciados no século passado, alcançando ampla expansão com a consolidação e a ubiquidade das mídias sociais.

2 Contexto

Um dos debates recorrentes sobre o conceito de desinformação é sua relação com informação e conhecimento. A desinformação é um tipo de informação? Ela produz conhecimento? É um fenômeno mental? Do contrário, o que ela é?

O tratamento conceitual sobre desinformação ocorre somente décadas após o seu principal conceito correlato, informação, ter uma centralidade na ciência do século XX. De acordo com Figueiredo, Ferrari e Lima (2023), é nessa época que a informação passa de mera palavra para um objeto de investigação de campos como a física, genética, linguística, computação, estatística, jornalismo, ciências cognitivas, filosofia, entre outros que "[...] trataram da informação de forma particular e a definiram de modo independente, revelando que, ao mesmo tempo que a noção de informação é fundamental para todas essas abordagens científicas, ela se diferencia a cada função que lhe é atribuída" (Figueiredo; Ferrari; Lima, 2023, p. 216).

A ideia contemporânea de informação está vinculada ao desenvolvimento da Teoria Matemática da Comunicação de Claude Shannon em 1948, posteriormente comentada por Warren Weaver, que transformou a linguagem formal de Shannon em uma explicação mais acessível, ao ponto de começar a ser entendida de fato como uma precursora Teoria da Informação, a exemplo do que

é proposto em Gleick (2016). Associada com a ideia de entropia importada do campo da Termodinâmica e acrescentada pela teoria cibernética desenvolvida por Norbert Wiener, a informação passou a ser compreendida como algo responsável por reduzir a desordem ou incerteza de um sistema, agindo como o negativo da entropia que subsiste como regra geral.

É pertinente mencionar que, nesse momento, as preocupações informacionais eram essencialmente sintáticas, direcionadas ao desenvolvimento científico e tecnológico de sistemas de comunicação e computação. Inserida em um diagrama de emissor-mensagem-receptor, a informação consistia no resultado de dados binários deslocados por sinais elétricos que buscavam efetivar o funcionamento desses sistemas.

Nesse contexto, a desinformação ainda estava formalmente desvinculada do conceito de informação e aparecia como objeto de outras áreas, por outros termos. Empregados atualmente pela literatura sob o rótulo "desinformação", os termos que surgem com mais evidência nessa época são, sobretudo, dois: propaganda e *hoax* (boato). O primeiro tem relação estreita com o vocabulário militar e geopolítico e é adotado para mencionar estratégias de influência da opinião pública por governos, como forma de manutenção dos regimes instaurados – com alguma frequência, por meios não democráticos. O segundo é atribuído à divulgação de histórias fabricadas, ação recorrente em tempos de jornalismo pré-industrial, como o caso do "boato da grande lua" (The Great Moon Hoax), publicado em seis artigos pelo jornal norte-americano The Sun, em 1835. Como pontua Dourado (2021), notícias veiculadas sem técnicas de apuração e deontologia jornalística eram algo comum no início do século XIX pela publicação de jornais populares conhecidos como *penny press*, parte da constituição do *yellow journalism* (traduzido em português como "imprensa marrom").

Ainda que os primeiros registros da palavra *misinformation* datem para o século XVI (Misinformation, 2026), a desinformação enquanto engano intencional (*disinformation*) aparece, a partir do conceito russo *dezinformatsiya*, somente no século XX (Acampa, 2024; Disinformation, 2026). Contudo, conforme relatam Freelon e Wells (2020), é importante destacar que poucas

contribuições teóricas sobre os conceitos foram feitas até o ano de 2017[1]. Segundo os autores, mesmo as investigações sobre a concepção de *misinformation*, que tomaram alguma centralidade de pesquisa no campo da psicologia cognitiva das décadas de 1970 e 1980, tiveram pouca expressividade de publicações. O que predominava na época, sobretudo nas pesquisas de comunicação, eram os estudos sobre *propaganda* inseridos em um contexto de crise da comunicação pública e desconfiança sobre a mídia jornalística.

Mesmo o cenário produzido pela nova Teoria da Informação – quando expandiu suas implicações epistemológicas e causou uma "virada informacional" que impactou diversas áreas do conhecimento (Gleick, 2016) – parece ter levado um considerável tempo para abordar a desinformação como uma questão fundamental nas reflexões sobre a natureza do conhecimento e sobre as relações mente/corpo e percepção/ação, que estavam sendo revistas por uma renovada perspectiva ontológica. Um dos primeiros esforços surge a partir da constituição do campo da Filosofia da Informação que, representado aqui pela abordagem de Floridi (2011), conceitua a informação enquanto substrato da realidade, presente no mundo de uma forma inexorável e constituinte de todos os organismos que habitam, nas palavras do autor, a "infosfera". É notável, assim, o tratamento que Floridi (1996) dá à desinformação – em um momento em que a World Wide Web ainda estava iniciando – enquanto um problema inerente a todo sistema de informação composto de fluxos e processos informacionais em grande escala. Isto é, a desinformação é uma consequência inevitável do funcionamento de qualquer sistema informacional complexo.

De forma geral, podemos considerar que, em relação ao objeto *informação*, um paradigma predominante até a primeira década dos anos 2000 foi aquele voltado ao desenvolvimento de tecnologias, à gestão de recursos documentais e, principalmente, à promoção do acesso à informação. Tornar a informação disponível ao maior número de pessoas possível parece um ponto de partida compartilhado, ainda que com execuções diferentes, entre os campos incluídos nas grandes áreas da Comunicação e Informação (Jornalismo, Publicidade, Biblioteconomia, Arquivologia e Ciência da Informação, entre outros).

Com a criação da World Wide Web, que trouxe a Internet ao grande público na década de 1990, uma série de movimentos teóricos e tecnológicos vislumbraram com otimismo as possibilidades permitidas pela rede mundial conectada por computadores. Instaurado este novo ambiente resultante da interconexão global, nomeado por Lévy (2010) de ciberespaço, o interesse social pelos fluxos de informação cada vez mais ágeis e participativos logo contribuiu com as transformações culturais e econômicas que encontraram no termo "sociedade da informação" uma espécie de *zeitgeist*, implicando considerável expectativa para com o compartilhamento de conhecimento sem fronteiras em favor de uma "inteligência coletiva" conectada e em pleno desenvolvimento.

De fato, observa-se um período historicamente singular de ampliação do acesso à informação, marcado pela digitalização e pela disponibilização *online* de conteúdos em larga escala. Esse movimento favoreceu a criação de repositórios, arquivos e acervos digitais por diferentes iniciativas, fortalecendo esforços institucionais de organização do conhecimento, gestão da informação e inclusão digital, em grande medida ancorados nos avanços produzidos pelos campos da Computação e da Ciência da Informação. Nesse período prolífico, foram instituídos marcos importantes que buscaram contribuir, por meio de trabalho cooperativo e voluntário, para uma cultura de democratização da informação e acesso livre ao conhecimento – a exemplo da Creative Commons, Wikipédia, Open Source Initiative, Internet Archive/Wayback Machine e do movimento do Acesso Aberto (Open Access), o qual permitiu a criação da SciELO, do Directory of Open Access Journals (DOAJ) e de outras bases de dados e bibliotecas digitais que seguem até hoje como referências na disponibilização irrestrita de publicações científicas. Cabe também destacar o crescimento que o Linux e o movimento do Software Livre tiveram nessa época.

No entanto, a partir de 2010, algo como um efeito colateral dessa explosão informacional começa a aparecer. As possibilidades da então *web* 2.0 junto com o surgimento das mídias sociais e a gradual inserção das instituições, organizações e empresas no ambiente digital, resultaram em um cotidiano de "hiperinformação" (Moretzsohn, 2017) que logo remodelou relações econômicas e sociais, das quais emergem as plataformas digitais geridas pelas *big techs* –

empresas que hoje controlam o fluxo majoritário de dados e serviços presentes no ambiente digital. Mais do que distribuir conteúdo, permitir a troca de informações e ampliar as conexões sociais, essas plataformas passaram a funcionar como estruturas que agem sobre a realidade e modulam comportamentos pela aplicação de sistemas algorítmicos dirigidos a maximizar o tempo de uso e o crescimento econômico dos seus ambientes. Conectadas aos sistemas social, político e econômico, operam como mercados especulativos que utilizam as interações como *commodities* e exploram financeiramente as atividades privadas de navegação na rede (Raulino, 2022), evidenciando um regime extrativo que coloca hoje sob discussão o colonialismo de dados praticado em países do sul global (Cassino; Souza; Silveira, 2021).

É relevante também considerar nesse contexto as mudanças causadas pelos canais de *partisan media*, primeiro no rádio e depois na TV a cabo, os quais segmentaram o mercado midiático e fomentaram a compreensão do público sobre informação enquanto um bem de consumo pessoal e de fortalecimento identitário, viabilizando a criação de nichos informacionais que mais tarde viriam a ser adotados pelas plataformas digitais (Freelon; Wells, 2020). O estímulo à indignação, a criação de controvérsias, a proeminência de conteúdos opinativos e um forte apelo ideológico são características fundantes desses canais, os quais passam a organizar perfis de consumidores que buscam informações alinhadas às suas crenças (Diaz Ruiz, 2025).

Após 2016, com a exposição das dinâmicas sociais que levaram à primeira eleição de Donald Trump à presidência dos Estados Unidos e à saída do Reino Unido da União Europeia (movimento nomeado de Brexit), os problemas no modelo otimizado da sociedade da informação foram evidenciados e incutiram ao vocabulário global as "*fake news*" e a "pós-verdade"; dois termos que seriam exaustivamente disseminados nos anos seguintes. O caso da Cambridge Analytica – que minerou, sem consentimento, uma quantidade massiva de dados de usuários do Facebook para realizar modelagem psicométrica e influenciar tomadas de decisão dos indivíduos com técnicas de *microtargetting* (Alvim, 2021) – foi outro momento definidor do cenário geopolítico, que revelou a lógica vigente por trás dos fluxos de informação que constituem o contexto contemporâneo. Esses

acontecimentos revelaram mudanças estruturais em curso que se intensificaram posteriormente, despertando a atenção das investigações científicas e da sociedade como um todo.

3 Conceito

Tal cenário provocou uma mudança de paradigma nos estudos e pesquisas sobre informação e comunicação, que passaram a direcionar seus esforços na compreensão do fenômeno da desinformação e formar gradativamente o que podemos chamar de uma incipiente Teoria da Desinformação associada à constituição do campo dos estudos sobre desinformação, consolidado pela literatura sob o termo *disinformation studies*.

Entretanto, a tarefa de conceituar rigorosamente "desinformação" parece demandar àquela empreendida em relação à sua raiz, informação, que produziu extensas investigações sobre o seu emprego nos diversos campos que a adotam como conceito (Capurro; Hjørland, 2007), reflexões epistemológicas sobre seu papel na constituição do conhecimento (Dretske, 2003) e aprofundamentos ontológicos sobre a sua característica de substrato da realidade (Floridi, 2011).

No sentido etimológico, como lembra Santaella (2023), informação deriva do latim *informare*, que significa "dar forma" a algo – uma ideia desenvolvida desde a filosofia grega, principalmente a partir de Aristóteles. Podemos assumir assim que "desinformação", com o seu prefixo indicando contrariedade do sentido da palavra a ele associado, em sua essência significaria "deformar", isto é, "corromper a forma" de algo.

Mas o que as definições encontradas sobre desinformação têm demonstrado? Em primeiro lugar, é preciso ressaltar a constante demanda por uma concepção pragmática geral que sustente as documentações e legislações que vêm sendo elaboradas nos últimos anos para lidar com as consequências políticas e sociais que o fenômeno tem causado. Nesse escopo, a desinformação tem sido entendida como uma questão de prejuízo ambiental (no sentido amplo da palavra) em uma sociedade cada vez mais digitalmente dependente em seus mecanismos de funcionamento.

Em documentos e publicações da Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO), da Organização Mundial de Saúde (OMS) e do Conselho da Europa (CoE), prevalece a ideia de que a desinformação é um ato intencional de induzir ao engano que tem finalidades geralmente políticas, ideológicas e/ou econômicas, operado por atores sociais e/ou grupos organizados. Enquanto parte de um cenário de "desordem informacional", a desinformação (*disinformation*) apresenta nuances e intersecções ao lado da informação incorreta/falsa (*misinformation*) e da má-informação/informação maliciosa[2] (*malinformation*) (Ireton; Posetti, 2019; Wardle; Derakshan, 2018; Wardle; Derakshan, 2023; WHO, 2024). Os documentos da UNESCO e do CoE ainda categorizam tipologias relacionadas aos níveis de intenção de dano, as quais incluem sátiras, equívocos cognitivos, descontextualizações, fraudes, vazamento de dados, assédio e discursos de ódio – o que, se por um lado evidencia a complexidade do comportamento informacional em rede, por outro acarreta problemas conceituais que comprometem a consistência dessas definições.

Já o Relatório de Riscos Globais, publicado anualmente pelo Fórum Econômico Mundial desde 2006, fornece uma definição que entende a desinformação como:

> [...] informações falsas recorrentes (intencionais ou não) amplamente divulgadas pelas redes de mídia, mudando a opinião pública de forma significativa para a desconfiança nos fatos e na autoridade. Inclui, mas não se limita a: conteúdo falso, impostor, manipulado e fabricado (World Economic Forum, 2026, p. 76, tradução nossa).

O relatório parece enfatizar as consequências e os meios de ação associados ao termo, mas deixa muito aberto o significado atribuído à noção de "autoridade" – palavra que, no presente contexto, poderia ser substituída por "instituições"; considerando, neste caso, as entidades que sustentam o modelo democrático, seja de governo, seja de mediação e produção do conhecimento.

Contudo, a ampla disseminação e variação terminológica que aconteceu nos últimos anos em diversas instâncias da esfera pública, causou distorções de sentido e utilizações inapropriadas que têm contribuído para um esvaziamento do significado efetivo daquilo que é nomeado. O caso das *fake news* é o mais evidente. Inicialmente usada para se referir aos *blogs* anônimos que imitavam o

layout[3] de portais de notícias para, com isso, transmitir credibilidade ao conteúdo veiculado, a expressão logo se tornou uma referência para falar sobre mentira de forma geral, ou mesmo para desqualificar discursos no campo retórico e repreender publicamente adversários políticos e veículos de comunicação[4]. Dourado (2021) aponta que o termo foi popularizado a partir das denúncias que o jornalista canadense Craig Silverman fez, a partir de 2014, sobre as *fake stories*: histórias sensacionalistas inventadas[5] para atrair tráfego pago para sites rudimentares, montados com o único intuito de gerar receita aos seus proprietários. Em razão da facilidade de criação, do custo praticamente nulo e das oportunidades de monetização, a estratégia foi rapidamente adotada por pessoas em diversas partes do mundo, sendo a Macedônia um dos casos mais emblemáticos[6].

Assim como a desinformação, registros históricos de uso do termo *fake news* remontam a várias décadas atrás. Na verdade, menções ao termo são encontradas em 1890, em comentários feitos dentro de jornais como Milwaukee (Wisconsin) Daily Journal, The Kearney Daily Hub e Cincinnati Commercial Tribune[7]. Mas da mesma forma que no caso da palavra "desinformação", esses usos eram esparsos e não se constituíam ainda enquanto referências a um objeto epistêmico. O próprio dicionário Oxford, apesar de apresentar registros do termo no século XIX, só criou a entrada do mesmo em sua lista de verbetes em 2019 (Fake news, 2026). É evidente que histórias sensacionalistas, teorias da conspiração e folhetins panfletários circulam desde muito tempo. Entretanto, consideramos que esses fenômenos constituem objetos analiticamente distintos, embora possam apresentar pontos de interseção em contextos específicos.

A dificuldade e variedade terminológicas refletem o crescimento da complexidade relativa ao próprio *habitat* digital. Se as *fake news* são resultantes de um modo de operação textual característico do início da Internet, a pós-verdade surge pela escalada do apelo retórico às emoções como código central dos discursos *online* e pelo início das conexões e transmissões audiovisuais em tempo real que permitiram *lives*, *streamings*, *podcasts* e outros formatos que revigoraram princípios de mídias como televisão, cinema e rádio, dentro do ciberespaço. Posteriormente, chegam os *deepfakes*, refletindo o aumento do poder

computacional associado ao desenvolvimento de técnicas de aprendizado profundo (*deep learning*) para a adulteração de vídeos com elevado grau de realismo. Com o advento da inteligência artificial generativa, outra etapa na criação de desinformação começou a ser efetivada, inserida no que propomos chamar de uma gradual *promptologia*, ou seja, em uma mudança de paradigma que insere o desenvolvimento humano em uma cultura de *prompts* – linhas de comando que alteram a forma de agir cognitivamente em meio aos fluxos de informação.

4 A desinformação enquanto questão sistêmica

A relação entre interesses econômicos das plataformas e desinformação converge na medida em que os mecanismos que produzem lucro, para os usuários produtores de conteúdo ou para as *big techs*, são os mesmos que estimulam a disseminação de desinformação. O apelo emocional e a exploração de heurísticas cognitivas, amplamente empregados nas estratégias de marketing midiático, são também as estratégias de disseminação da desinformação, resultando em um ecossistema que se retroalimenta. A desinformação passa a ser mobilizada por diversos atores enquanto um modelo de negócio. Mesmo quando não é produzida ou estimulada diretamente pelas plataformas digitais ou por agentes específicos da produção de conteúdo, a desinformação continua a gerar benefícios indiretos se o seu status existencial é relativizado ou negligenciado, uma vez que ela impulsiona o fluxo de interações que sustenta a economia de extração e comercialização de dados. Isso significa, conforme análise feita por Diaz Ruiz (2025), que no mercado da mídia digital, a desinformação não é um problema, mas sim um resultado esperado.

Durante algum tempo, houve certa convergência na literatura para as definições adotadas pelos documentos das instituições e entidades mencionados anteriormente. Isto é, a desinformação (no sentido da *disinformation*) tende a ser entendida como o ato intencional de produzir engano utilizando algum conteúdo semântico, considerando que o objetivo desse ato pode incluir também a replicação do conteúdo pelos indivíduos afetados (*misinformation*), ampliando o alcance do efeito causado. No entanto, podemos considerar que se trata de uma

definição dicionarizada, como aponta Mona Simion (2024), que pouco avança epistemologicamente e se mostra insuficiente diante de diagnósticos mais profundos sobre o fenômeno.

Estudos como os de Recuero (2024, 2025), Diaz Ruiz (2025), Amaral e Cardoso (2025) e Schneider (2022) têm enfatizado essa insuficiência conceitual diante de uma questão que abrange um ecossistema dinâmico em constante crescimento. Recuero (2025, p. 2, tradução nossa) também questiona até que ponto é possível afirmar a presença de intenção dentro de um fluxo informacional, considerando que mesmo proposições verdadeiras podem ser enquadradas de forma a produzir crenças falsas e que "[...] "intenção", em muitos casos de disseminação de conteúdo, é difícil de determinar, e muitas vezes torna-se impraticável aplicar rigorosamente a distinção entre *mis-disinformation* em pesquisas empíricas ou práticas de moderação de conteúdo".

Mas talvez a principal questão relacionada à desinformação sistêmica seja o desenvolvimento das tecnologias cognitivas, representado no atual estágio pela expansão da inteligência artificial generativa e pela emergência da inteligência artificial agêntica, que se distingue pela capacidade de planejar, tomar decisões e executar ações com elevados níveis de autonomia, ampliando significativamente o potencial de intervenção dos sistemas computacionais nos fluxos informacionais (Schroeder *et al.*, 2026). Podemos considerar que essas tecnologias agem, seguindo o que é proposto por Floridi (2014), enquanto organismos informacionais, pois afetam, com crescente liberdade, o ciclo informacional de criação, coleta, armazenamento, processamento, distribuição, consumo e descarte. A esse respeito, designamos como "tecnologias cognitivas" por entendermos que estamos diante de um estágio de desenvolvimento que sucede as tecnologias da informação e da comunicação, caracterizado pelo papel ativo dos algoritmos na construção de sentido nos fluxos informacionais e pelo desenvolvimento de modelos computacionais de inteligência capazes de expandir as capacidades cognitivas humanas.

A inteligência artificial (IA) generativa alcançou ampla visibilidade pública com o lançamento do ChatGPT pela empresa OpenAI, em 2022. Fundamentada em técnicas de *deep learning*, essa tecnologia depende do

processamento de grandes volumes de dados e de ciclos contínuos de aperfeiçoamento conduzidos pelas empresas desenvolvedoras, frequentemente apoiados nas interações realizadas pelos usuários em suas plataformas e serviços.

A disseminação pública da IA generativa foi primeiro repercutida no horizonte textual, resultado da implementação de *chatbots* para fins diversos, mas logo ampliada ao campo audiovisual, onde permitiu o aperfeiçoamento técnico e a simplificação dos processos na criação de *deepfakes*. Esse aprimoramento tornou possível a elaboração de vídeos falsos sem o emprego de materiais pré-existentes (requisito até então indispensável), inaugurando uma nova forma de desinformação: a desinformação sintética. Isto é, com o avanço tecnológico – materializado pelos modelos presentes, por exemplo, no Kling AI (da empresa Kuaishou), Gen (da Runway), Veo e Nano Banana (Google) – instaurou-se um cenário inédito de produção de conteúdos audiovisuais inteiramente feito por *prompts*. Outro fator dentro deste contexto é a possibilidade de captura de movimentos em tempo real pela câmera do celular ou computador para contribuir na efetividade de crimes digitais, como golpes e fraudes, por meio da construção de identidades falsas em videochamadas e transmissões ao vivo.

Portanto, no que tange ao emprego de tecnologias cognitivas para o refinamento na produção de desinformação em operações sistêmicas, emergem ao menos duas questões adicionais. A primeira é um maior desenvolvimento das automações utilizadas nos fluxos de desinformação, como a criação de perfis *fakes* e *bots* com funções de compartilhamento, *trolling* e outras ações que contribuem para um conteúdo ou assunto receber visibilidade e repercussão. Nesse sentido, o risco inclui um aprofundamento da técnica conhecida como *astroturfing*, a criação artificial de tendências por uma planejada e encoberta coordenação automatizada de múltiplas contas nas plataformas digitais (Salaverría; Cardoso, 2023). Por meio do *astroturfing*, é possível mobilizar ciclos organizados e programados de promoção de desinformação que são percebidos pelos algoritmos das plataformas e pela opinião pública como manifestações espontâneas e orgânicas, criando a falsa impressão de uma convergência natural de interesses em torno de determinado tema. Dessa forma, colaboram para o que, segundo Recuero (2025, p. 2, tradução nossa), os estudos mais recentes sobre

desinformação vêm chamando de "operações de influência", "[...] tentativas coordenadas e estratégicas de influenciar a opinião pública em todos os ecossistemas midiáticos – indo além de simples atos de desinformação".

Em uma pesquisa publicada na revista *Science* e conduzida por 21 pesquisadores de sete países, foram relatadas as primeiras evidências de um avanço significativo nas operações de comportamento coordenado inautêntico (*coordinated inauthentic behavior*). As operações estão associadas às chamadas *AI swarms*, resultado da integração entre Large Language Models (LLMs) e arquiteturas computacionais multiagentes. Essa combinação possibilita a criação de sistemas nos quais múltiplas IAs atuam de forma coordenada em torno de um objetivo comum, mas executando tarefas distintas e adaptando seu comportamento, em tempo real, às interações humanas e às vulnerabilidades das plataformas digitais (Schroeder *et al*., 2026). Segundo os autores, o cenário coloca as operações de influência sintética em um nível sem precedentes, fomentando um ecossistema composto por fluxos ininterruptos de conteúdos e interações artificialmente fabricados; o que fortalece a possibilidade de consolidação de ambientes informacionais amplamente permeados por desinformação. Embora o estudo identifique um aumento do ceticismo público diante de conteúdos não verificados, o risco apontado pelos autores é que a dinâmica de automação adaptativa desses sistemas seja capaz de contornar, de forma crescente, os mecanismos de verificação e detecção.

A segunda questão relacionada ao papel das tecnologias cognitivas na produção de desinformação diz respeito à contaminação dos dados de entrada em sistemas de inteligência artificial. Até que ponto os conjuntos de dados empregados no treinamento dessas tecnologias não contêm informações falsas, distorcidas ou enviesadas que comprometem a qualidade e a confiabilidade dos conteúdos posteriormente gerados?[8] Schroeder *et al*. (2026) apontam a existência de redes feitas com o propósito exclusivo de contaminar dados de treinamento dos modelos de IA, a exemplo da russa Pravda. Segundos os pesquisadores:

> Essas redes parecem ter sido criadas especificamente para serem consumidas por máquinas. A duplicação de artigos em centenas de domínios, interfaces de usuário de baixa qualidade e baixo tráfego humano indicam que seu público-alvo principal são os rastreadores web que alimentam os modelos de linguagem de grande escala (LLMs) (Schroeder *et al*., 2026, p. 3, tradução nossa).

Associado com as *AI swarms*, o cenário pode aprofundar os riscos epistêmicos: "Esses enxames podem manipular o substrato epistêmico no qual se basearão as futuras deliberações e ferramentas de IA, prejudicando o debate público informado do qual depende a democracia" (Schroeder *et al*., 2026, p. 3, tradução nossa). Ainda nesse âmbito, é fundamental examinar as estruturas lógicas e os critérios de modelagem que orientam o funcionamento dos algoritmos, pois eles podem incorporar vieses, preconceitos e falhas éticas aos processos computacionais que condicionam os resultados gerados pelo sistema, conforme relatado pelas análises do Fórum Econômico Mundial (World Economic Forum, 2026) e por pesquisas como a de Silva (2024).

Um último elemento sobre essa questão é a emergência das chamadas "alucinações", fenômeno em que o sistema gera respostas que não encontram respaldo nos dados utilizados em seu treinamento nem podem ser justificadas por inferências válidas, resultando na apresentação de informações falsas, imprecisas ou inexistentes. Na contramão do entendimento usual desse processo enquanto resultado de um "*bug*" ou erro eventual de processamento, Floridi, Jia e Tohme (2025) argumentam que, na verdade, se trata de uma característica estrutural e intrínseca desses modelos, que não pode ser tecnicamente erradicada. É como se as LLMs, então, fossem inevitavelmente falíveis por *design*, por "natureza". Isso acontece porque elas manipulam símbolos por processos estatísticos que não estão conectados com a referência genuína da experiência. Os autores exemplificam de forma simples: "O LLM aprende que os humanos dizem "Paris é a capital da França" em contextos específicos, e não que Paris é a capital da França" (Floridi; Jia; Tohme, 2025, p. 15, tradução nossa). Ou seja, o que a IA processa:

> [...] são padrões na linguagem humana sobre o mundo, não padrões no mundo em si. O LLM detecta regularidades de segunda ordem — padrões na forma como os humanos descrevem padrões — sem acessar as regularidades de primeira ordem que fundamentam as descrições humanas (Floridi; Jia; Tohme, 2025, p. 15, tradução nossa).

Assim, embora as LLMs sejam capazes de produzir resultados sofisticados, sua operação permanece fundamentada em processos estatísticos de manipulação de símbolos que são compreendidos em seu significado somente por agentes humanos. As alucinações, portanto, são falhas inatas de implementação justamente porque a semântica produzida não é apreendida fenomenologicamente, como acontece com a experiência humana. Padrões estatisticamente plausíveis nem sempre correspondem a afirmações verdadeiras e, segundo Floridi, Jia e Tohme (2025), padrões podem ser produzidos por dados de treinamento que contém contextos estatísticos incomuns ou através da ativação de associações errôneas pelos parâmetros do *prompt*.

Floridi, Jia e Tohme (2025) ainda sugerem que mesmo que sejam integrados mecanismos sensório-motores que façam a IA mapear ambientes como parte da formulação de informações, essa relação continua não sendo genuína pois ainda está guiada pelo *grounding* humano presente no *corpus* de treinamento. Isto é:

> O LLM não estabelece suas próprias conexões causais e informacionais [...]. Assim, quaisquer relações referenciais obtidas surgem de forma derivada, por meio da mediação de conteúdo de origem humana, e não por meio do próprio envolvimento do LLM com o mundo (Floridi; Jia; Tohme, 2025, p. 17, tradução nossa).

Vinculado ao fenômeno das alucinações encontra-se também o processo de *sycophancy*, compreendido como um viés de concordância pelo qual *chatbots* tendem a validar ou reforçar as alegações dos usuários, mesmo quando estas são incorretas. Esse comportamento é atribuído aos mecanismos de aprendizado baseados em *feedback* humano, que tendem a favorecer respostas percebidas como úteis, satisfatórias ou agradáveis. Nesse sentido, o modelo pode gerar alucinações para corroborar com as crenças e expectativas do usuário – o que, conforme relatado por Chandra *et al*. (2026), pode contribuir com processos delirantes e intensificar sintomas psicóticos em indivíduos.

O panorama delineado indica que a desinformação constitui, na atualidade, um fenômeno sistêmico, instrumentalizado por atores envolvidos em disputas políticas e econômicas e favorecido pela arquitetura e pelos modelos de negócio das plataformas digitais. Estas, por sua vez, acumulam poder econômico precisamente em razão da intensificação da circulação, do engajamento e da monetização de tudo que transita em seus ambientes. Recuero (2024) argumenta que a desinformação só pode ser efetivamente enfrentada se compreendida enquanto um sistema desinformativo, cujo efeito é o crescimento da entropia informacional em níveis capazes de gerar desordem e instabilidade nos sistemas sociais. Mas como desestabilizar este sistema?

O avanço das tecnologias cognitivas e a progressiva imersão da subjetividade humana nos ambientes digitais produzem uma configuração inédita da experiência social, na qual os processos de significação são produzidos por camadas de representação cada vez mais dissociadas da materialidade factual. Nesse sentido, é instigante a provocação de Guarda, Ohlson e Romanini (2018) ao propor a iminência de uma "pós-realidade". Segundo os autores, mecanismos de filtros-bolha e câmaras de eco inserem indivíduos e grupos em universos semióticos herméticos, onde frequentemente fatos e evidências são rejeitados. Como resultado, ocorre a construção de representações mentais que podemos remeter, por analogia, às próprias "alucinações" das LLMs.

Essa perspectiva abre caminho para investigações semióticas sobre a desinformação, que consideramos urgentes para compreender o fenômeno em suas dimensões conceituais, cognitivas e sistêmicas. Contudo, considerando o escopo e os objetivos deste texto, faremos apenas um breve comentário.

Signos são, na perspectiva semiótica de Charles Sanders Peirce, mediações representacionais entre objetos e interpretantes, produzindo efeitos de sentido que possibilitam os processos de significação (CP 2.227-228)[9]. Nos processos habituais de produção de informação, busca-se o refinamento contínuo da relação entre signo e objeto, de modo que os efeitos interpretantes produzidos correspondam, tanto quanto possível, às características do objeto representado. Nos fluxos de desinformação, em contraste, promovem-se operações de deslocamento ou ruptura dessa relação, pelas quais signos socialmente

estabilizados passam a produzir interpretantes dissociados de seus objetos de referência e orientados por interesses políticos, ideológicos ou econômicos dos agentes envolvidos em sua disseminação.

Esses efeitos, amplificados por estratégias de modulação cognitiva fomentadas pela economia da atenção vigente no ambiente digital, induzem a sedimentação de crenças e disposições interpretativas, afetando amplos contingentes de indivíduos em curtos intervalos de tempo.

É preciso considerar, entretanto, que a progressiva digitalização da experiência social produziu importantes reconfigurações semióticas de conceitos fundamentais, como realidade, identidade e verdade. Além disso, os ambientes digitais deram origem a novos signos e formas de comunicação, entre os quais se destacam os *memes*, que frequentemente atuam como veículos sutis de circulação e disseminação da desinformação. O cenário corrobora a pertinência da noção de realidade *onlife* proposta por Floridi (2014), segundo a qual fronteiras outrora entendidas como categóricas – como *online* e *offline*, natural e artificial, biológico e tecnológico – tornam-se progressivamente difusas, configurando novas formas de experiência, interação e produção de sentido.

O ecossistema da desinformação é, sobretudo, um conjunto de estratégias e processos que se valem da tecnologia para corromper signos socialmente compartilhados, com o objetivo de produzir crenças e orientar hábitos de conduta que favoreçam os agentes que o sustentam.

5 Considerações finais

Investigar o fenômeno da desinformação é um desafio analítico exigente pois envolve um dos fatores que mais mobilizam o tecido social na atualidade e está profundamente interligado às dinâmicas socioeconômicas e políticas que estruturam o mundo contemporâneo. Sobretudo, a desinformação tornou-se um modo de ação amplamente disseminado pelas plataformas digitais e mobilizado por uma parcela significativa de atores sociais e instituições que disputam a opinião pública e o controle dos mecanismos de poder, cada vez mais globalizados e difusos.

Considerando a perspectiva da filosofia da informação, podemos questionar em que medida informação e desinformação, em sentido ontológico, se constituem de maneira similar. A informação pode ser compreendida como um princípio metafísico que viabiliza processos evolutivos, enquanto a desinformação está associada ao aumento da entropia e à degradação de tais processos. A informação apresenta-se, assim, como uma variável natural, imanente ao mundo, pois consiste na essência do funcionamento de mecanismos vitais de todos os tipos. Seria possível sustentar o mesmo estatuto ontológico para a desinformação?

Compreendemos que a desinformação pode ser definida, de modo geral, como uma falha (*misinformation*) ou como uma distorção deliberada (*disinformation*) nos processos de captação, processamento e transmissão de informação por agentes cognitivos. No caso da distorção deliberada, a literatura frequentemente a compreende como uma ação orientada a influenciar outro agente. Porém, é possível também considerar a dimensão de processos de autoengano epistêmico, sobretudo quando certos argumentos deslocam a *misinformation* de um simples equívoco para uma aceitação ativa de crenças falsas (Freelon; Wells, 2020). No entanto, ao assumir o princípio falibilista inerente à ciência, reconhece-se que a adesão a crenças falsas, mesmo que não intencionalmente, constitui um elemento recorrente na história do conhecimento e, por extensão, na própria história humana. Isso abre espaço para reflexões sobre o eventual estatuto ontológico da desinformação.

Com o advento das tecnologias cognitivas, o cenário tende a assumir contornos mais complexos e potencialmente mais preocupantes. As pesquisas sobre desinformação passam a considerar um novo estágio, no qual a produção de conteúdo envolve agentes artificiais com diferentes graus de autonomia, cujos mecanismos de funcionamento não são plenamente transparentes ao escrutínio público, e que podem ser instrumentalizados por atores humanos em contextos de disputa informacional e de *information warfare*, orientados por estratégias de influência e governança.

Por fim, este texto, orientado por uma reflexão epistemológica sobre os efeitos observados no cotidiano dos fluxos informacionais contemporâneos,

buscou contribuir para o avanço dos estudos sobre desinformação, em especial para o que se pode delinear como uma possível teoria da desinformação, situada entre os campos da Ciência da Informação e da Comunicação, mas com caráter multidisciplinar.

Financiamento

Pesquisa realizada com apoio de bolsa da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Código de Financiamento 001.

Referências

ACAMPA, Suania. **From dezinformatsiya to disinformation**: a critical analysis of strategies and effect on the digital public sphere. Cham: Springer, 2024.

ALVIM, Marcelo Hamdan. Uma leitura pragmaticista e comunicacional sobre a exploração de dados e mediações algorítmicas: o caso Cambridge Analytica. **Semeiosis**: **Semiótica e Transdisciplinaridade em Revista**, São Paulo, v. 9, n. 1, p. 104-118, 2021.

AMARAL, Gustavo Rick; CARDOSO, Tarcísio. Uma abordagem semiótico-cognitiva do fenômeno da desinformação. **Esferas**, Brasília, v. 1, n. 32, p. 1-25, 2025. Disponível em: https://doi.org/10.31501/esf.v1i32.15505. Acesso em: 09 jan. 2026.

CAPURRO, Rafael; HJØRLAND, Birger. O conceito de informação. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 12, n. 1, p. 148-207, 2007.

CASSINO, João Francisco; SOUZA, Joyce; SILVEIRA, Sérgio Amadeu (org.). **Colonialismo de dados**: como opera a trincheira algorítmica na guerra neoliberal. São Paulo: Autonomia Literária, 2021.

CHANDRA, Kartik *et al*. Sycophantic chatbots cause delusional spiraling, even in ideal Bayesians. **arXiv**, Ithaca, fev. 2026. Preprint. Disponível em: https://doi.org/10.48550/arXiv.2602.19141. Acesso em: 10 abr. 2026.

DIAZ RUIZ, Carlos. Disinformation on digital media platforms: a market-shaping approach. **New Media & Society**, London, v. 27, n. 4, p. 2188-2211,

2025. Disponível em: https://doi.org/10.1177/14614448231207644. Acesso em: 28 jan. 2026.

DISINFORMATION. *In*: OXFORD English Dictionary. Oxford: Oxford University Press, c2026.

DOURADO, Tatiana. **Fake news**: quando mentiras viram fatos políticos. Porto Alegre: Zouk, 2021.

DRETSKE, Fred. **Knowledge and the flow of information**. 74. ed. rev. Stanford: Center for the Study of Language and Information, 2003.

FAKE news. *In*: OXFORD English Dictionary. Oxford: Oxford University Press, c2026.

FIGUEIREDO, Suely; FERRARI, Luís Carlos; LIMA, Davino. **Desinformação, ingrediente da desordem**: um resgate crítico de fake news. Curitiba: Appris, 2023.

FLORIDI, Luciano. Brave.Net.World: the Internet as a disinformation superhighway?. **The Electronic Library**, Leeds, v. 14, n. 6, p. 509-514, 1996. Disponível em: https://doi.org/10.1108/eb045517. Acesso em: 08 jan. 2026.

FLORIDI, Luciano. **The 4th revolution**: how the infosphere is reshaping human reality. Oxford: Oxford University Press, 2014.

FLORIDI, Luciano. **The philosophy of information**. Oxford: Oxford University Press, 2011.

FLORIDI, Luciano; JIA, Yiyang; TOHME, Fernando. A categorical analysis of Large Language Models and why LLMs circumvent the symbol grounding problem. **arXiv**, Ithaca, dez. 2025. Preprint. Disponível em: http://dx.doi.org/10.2139/ssrn.5894082. Acesso em: 05 fev. 2026.

FREELON, Deen; WELLS, Chris. Disinformation as political communication. **Political Communication**, London, v. 37, n. 2, p. 145-156, 2020. Disponível em: https://doi.org/10.1080/10584609.2020.1723755. Acesso em: 20 jan. 2026.

GLEICK, James. **A informação**: uma história, uma teoria, uma enxurrada. São Paulo: Companhia das Letras, 2016.

GUARDA, Rebeka Figueiredo; OHLSON, Marcia Pinheiro; ROMANINI, Vinicius. Disinformation, dystopia and post-reality in social media: a semiotic-

cognitive perspective. **Education for Information**, Thousand Oaks, v. 34, n. 3, p. 185-197, 2018. Disponível em: https://doi.org/10.3233/EFI-180209. Acesso em: 10 dez. 2025.

HUGHES, Heather C.; WAISMEL-MANOR, Israel. The Macedonian fake news industry and the 2016 US election. **Political Science & Politics**, Cambridge, v. 54, n. 1, p. 19-23, 2021. Disponível em: https://doi.org/10.1017/S1049096520000992. Acesso em: 06 abr. 2026.

IRETON, Cherilyn; POSETTI, Julie. **Jornalismo, fake news & desinformação**: manual para educação e treinamento em jornalismo. Paris: UNESCO, 2019.

JAMIESON, Amber. 'You are fake news': Trump attacks CNN and BuzzFeed at press conference. **The Guardian**, London, 11 jan. 2017.

LÉVY, Pierre. **Cibercultura**. 3. ed. São Paulo: Edições 34, 2010.

MEHDIZADEH, Aliakbar; HILBERT, Martin. Epistemic substitution: how Grokipedia's AI-generated encyclopedia restructures authority. **arXiv**, Ithaca, dez. 2025. Preprint. Disponível em: https://doi.org/10.48550/arXiv.2512.03337. Acesso em: 11 abr. 2026.

MERRIAM-WEBSTER. **The real story of 'fake news'**. Springfield: Merriam-Webster, c2026.

MISINFORMATION. *In*: OXFORD English Dictionary. Oxford: Oxford University Press, c2026.

MORETZSOHN, Sylvia Debossan. "Uma legião de imbecis": hiperinformação, alienação e o fetichismo da tecnologia libertária. **Liinc em Revista**, Rio de Janeiro, v. 13, n. 2, p. 294-306, 2017. Disponível em: https://doi.org/10.18617/liinc.v13i2.4088. Acesso em: 06 mar. 2026.

MURPHY, Gillian *et al*. What do we study when we study misinformation?: a scoping review of experimental research (2016-2022). **Harvard Kennedy School Misinformation Review**, Cambridge, v. 4, n. 6, p. 1-57, 2023. Disponível em: https://doi.org/10.37016/mr-2020-130. Acesso em: 06 mar. 2026.

RAULINO, Gabriela. Capital e trabalho nas plataformas sociodigitais. *In*: DANTAS, Marcos *et al*. **O valor da informação**: de como o capital se apropria

do trabalho social na era do espetáculo e da internet. São Paulo: Boitempo, 2022. p. 145-212.

RECUERO, Raquel. **A rede da desinformação**: sistemas, estruturas e dinâmicas nas plataformas de mídias sociais. Porto Alegre: Sulina, 2024.

RECUERO, Raquel. A systemic framework for disinformation on social media platforms. **Platforms & Society**, Thousand Oaks, v. 2, p. 1-12, 2025. Disponível em: https://doi.org/10.1177/29768624251367199. Acesso em: 04 mar. 2026.

SALAVERRÍA, Ramón; CARDOSO, Gustavo. Future of disinformation studies: emerging research fields. **Profesional de la información**, Barcelona, v. 32, n. 5, p. 1-7, 2023. Disponível em: https://doi.org/10.3145/epi.2023.sep.25. Acesso em: 28 fev. 2026.

SANTAELLA, Lucia. Definir desinformação é preciso. *In*: SANTAELLA, Lucia (org.). **Flagelos da desinformação**. São Paulo: EDUC, 2023. p. 25-54.

SCHNEIDER, Marco. **A era da desinformação**: pós-verdade, fake news e outras armadilhas. Rio de Janeiro: Garamond, 2022.

SCHROEDER *et al*. How malicious AI swarms can threaten democracy. **Science**, Washington, v. 391, n. 6783, p. 354-357, 2026. Disponível em: https://doi.org/10.1126/science.adz1697. Acesso em: 23 mar. 2026.

SILVA, Tarcízio. **Racismo algorítmico**: inteligência artificial e discriminação nas redes digitais. São Paulo: Sesc, 2024.

SIMION, Mona. Knowledge and disinformation. **Episteme**, Cambridge, v. 21, n. 4, p. 1208-1219, 2024. Disponível em: https://doi.org/10.1017/epi.2023.25. Acesso em: 09 jan. 2025.

TRIEDMAN, Harold; MANTZARLIS, Alexios. What did Elon change? A comprehensive analysis of Grokipedia. **arXiv**, Ithaca, nov. 2025. Preprint. Disponível em: https://doi.org/10.48550/arXiv.2511.09685. Acesso em: 10 abr. 2026.

VAN DIJCK, Jose. Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. **Surveillance & Society**, Kingston, v. 12, n. 2, p. 197-208, 2014. Disponível em: https://doi.org/10.24908/ss.v12i2.4776. Acesso em: 10 abr. 2026.

YASSERI, Taha; MOHAMMADI, Saeedeh. How similar are Grokipedia and Wikipedia? A multi-dimensional textual and structural comparison. **arXiv**, Ithaca, jan. 2026. Preprint. Disponível em: https://doi.org/10.48550/arXiv.2510.26899. Acesso em: 10 abr. 2026.

WARDLE, Claire; DERAKHSHAN, Hossein. **Desordem informacional**: para um quadro interdisciplinar de investigação e elaboração de políticas públicas. Campinas: Unicamp, 2023.

WARDLE, Claire; DERAKSHAN, Hossein. **Information disorder**: toward an interdisciplinary framework for research and policy making. 2. ed. rev. Strasbourg: Council of Europe, 2018.

WHO. **Disinformation and public health**: questions and answers. Genebra: World Health Organization, 6 fev. 2024.

WORLD ECONOMIC FORUM. **The global risks report 2026**: 21st edition: insight report. Genebra: World Economic Forum, 2026.

Disinformation as a systemic problem: contributions to an epistemological discussion on the phenomenon

Abstract: This paper aims to contribute to the understanding of contemporary disinformation as a systemic phenomenon, as highlighted by recent research in the field of disinformation studies. Characterized by the environment of digital platforms, we argue that this phenomenon is intensified by the development of cognitive technologies, currently represented by generative artificial intelligence and the emerging agentic artificial intelligence. To support this analysis, we relate the terminological origin of disinformation to the context of the informational turn in scientific and philosophical thought, in addition to discussing contemporary conceptions of the topic present in both scientific literature and documents from international organizations. We argue that disinformation transcends its etymological definition and has established itself as a political and economic tool, exploited within dynamics structured by big tech companies in the realm of cyberspace. Finally, we highlight the urgency of reevaluating informational and communicational relationships in the face of the risk of a "post-reality" and defend the relevance of the semiotic perspective for understanding the processes of meaning production that traverse the flows of information and disinformation.

Keywords: disinformation; digital platforms; artificial intelligence; fake news; philosophy of information

Declaração de autoria

Concepção e elaboração do estudo: Leonardo Ripoll

Coleta de dados: Leonardo Ripoll

Análise e interpretação de dados: Leonardo Ripoll

Redação: Leonardo Ripoll

Revisão crítica do manuscrito: Leonardo Ripoll

Declaração de disponibilidade de dados

Não se aplica.

Autoria para correspondência

Leonardo Ripoll

leonardo.ripoll@usp.br

Editor-chefe

Thiago Henrique Bragato Barros

Como citar

RIPOLL, Leonardo. A desinformação como um problema sistêmico: contribuições para uma reflexão epistemológica sobre o fenômeno. **Em Questão**, Porto Alegre, v. 32, e-150997, 2026. DOI: https://doi.org/10.1590/1808-5245.33.150997

Recebido: 16/10/2025

Aceito: 21/08/2026

Copyright (c) 2026 Leonardo Ripoll. Este trabalho está licenciado sob uma licença Creative Commons Attribuição 4.0 Internactional. Autores mantêm os direitos autorais e concedem à revista o direito de primeira publicação, com o trabalho licenciado sob a Licença Creative Commons Attribution (CC BY 4.0), que permite o compartilhamento do trabalho com reconhecimento da autoria. Os artigos são de acesso aberto e uso gratuito.

[1] Tal fato pode ser constatado ao se realizar uma busca simples em qualquer base de dados acadêmica. No Portal de Periódicos da CAPES, por exemplo, 1402 documentos foram encontrados, em abril de 2026, pelo uso da palavra *disinformation* somente no campo "título". Destes, 1.341 itens (95,6%) são textos criados a partir de 2017. Usando a palavra *misinformation*, o resultado traz 3.382 documentos, dos quais 3.022 (89,3%) foram publicados a partir de 2017. Na base Scopus, os números são ainda maiores. Ao todo, foram 3.023 documentos recuperados pelo uso de "*disinformation*" no campo de título e 6.204 documentos listados para a palavra *misinformation*. No caso de disinformation, 2.866 (94,8%) surgiram a partir de 2017. Para *misinformation*, a correspondência é de 5.511 (88,8%) itens.

[2] Há uma diferença de tradução sobre os termos *misinformation* e *malinformatio*n entre a versão em português do manual da UNESCO e a tradução do relatório do Conselho da Europa, publicada pelo Centro de Lógica e Epistemologia (CLE) da Unicamp. Embora o documento da UNESCO (Ireton; Posetti, 2019) também se baseie no esquema conceitual apresentado por Wardle e Derakshan (2018), ele traduz *misinformation* como informação incorreta e *malinformation* como má-informação. Já os termos escolhidos pela tradução do relatório do conselho europeu (Wardle; Derakshan, 2023) foram, respectivamente, "informação falsa" e "informação maliciosa". Essas divergências acabam contribuindo para uma confusão semântica e conceitual dos termos, uma vez que trazem conotações diferentes.

[3] Entre essas imitações, são frequentes os casos de páginas que reproduzem logotipos e adotam nomes muito semelhantes aos de veículos jornalísticos amplamente reconhecidos. É por isso que Dourado (2021) chega a definir *fake news* como "contrafações noticiosas". A autora destaca a recorrência de termos como "Daily", "News", "Diário" e "Notícias" no nome desses sites, observando ainda que, em muitos casos, as estratégias de imitação incluem a adoção de endereços eletrônicos parecidos aos de portais consolidados.

[4] Um exemplo notável nesse sentido foi o momento em que Donald Trump, durante uma coletiva de imprensa em seu primeiro mandato como presidente dos Estados Unidos em 2017, acusou o canal de notícias CNN de ser *fake news* (Jamieson, 2017). Como argumenta Dourado (2021, p. 10), no campo político o termo se tornou "[...] um escape conveniente às polêmicas cotidianas, às acusações de adversários, às notícias indesejadas e a processos de responsabilização".

[5] Diferentemente das sátiras e paródias criadas por jornais como o *The Onion* ou os portais brasileiros *Sensacionalista* e *The Piauí Herald*, essas histórias eram promovidas enquanto notícias factuais. No entanto, é preciso considerar também que as sátiras e paródias, caso não sejam percebidas como tais, podem gerar *misinformation*.

[6] A cidade de Veles, na Macedônia do Norte, foi matéria de diversas investigações jornalísticas que a relacionam como um centro ativo de produção de *fake news* durante as eleições dos Estados Unidos em 2016. Um estudo sobre o caso foi publicado por Hughes e Waismel-Manor (2021).

[7] É o que registra tanto o Dicionário Oxford (Fake news, 2026) quanto o Merriam-Webster (2026). O último complementa que o substantivo composto *false news* aparece muito anteriormente, com registros encontrados no século XVI. Dourado (2021) também menciona a expressão *fake journalism*, usada em 1898 por de J. B. Montgomery-M'Govern ao se referir às crescentes publicações da imprensa marrom da época.

[8] O lançamento da Grokipedia pela xAI, empresa de Elon Musk, no final de 2025, representa um caso emblemático de reconfiguração epistemológica do acesso à informação, ao alterar os mecanismos pelos quais conteúdos são produzidos, organizados, interpretados e legitimados no ambiente digital. Proposta enquanto uma alternativa à Wikipédia, a nova enciclopédia *online*

tem todo o seu conteúdo gerado pela inteligência artificial Grok. Alguns estudos já apontam diversos problemas na criação dos artigos, como a ampliação do fluxo retórico em detrimento da diminuição de citações e referências, o plágio seletivo de textos da Wikipédia e o uso de fontes pouco confiáveis e alheias ao consenso científico (Yasseri; Mohammadi, 2026; Triedman; Mantzarlis, 2025; Mehdizadeh; Hilbert, 2025). Outra questão, sugerida pelos dados analisados por Yasseri e Mohammadi (2026), é o enviesamento dos conteúdos sob os valores políticos e pessoais de Elon Musk. Ou seja, se na Wikipédia as entradas são feitas coletivamente e balizadas pelos voluntários do projeto (respeitando as diretrizes de uso da Wikimedia Foundation), na Grokipedia ocorre uma produção centralizada de conteúdo, feita por uma inteligência artificial orientada a mimetizar cognitivamente o seu proprietário, acessando fontes pré-determinadas e reenquadrando o conteúdo sob a perspectiva ideológica de Elon Musk. Para os autores, "A natureza aberta da Wikipédia torna os vieses visíveis e contestáveis por meio de edições, disputas e deliberações [...]. A Grokipedia substitui esse processo por uma autoria opaca e automatizada, incorporando possíveis vícios ao comportamento do modelo, em vez de expô-los ao escrutínio. Apesar de seu objetivo corretivo declarado, a Grokipedia funciona menos como uma alternativa epistêmica e mais como uma reconfiguração da Wikipédia mediada por IA, trocando a responsabilidade coletiva pela autoridade computacional" (Yasseri; Mohammadi, 2026, p. 16, tradução nossa).

[9] Seguimos a convenção adotada mundialmente nos estudos peirceanos, uma vez que a maior parte dos textos de Peirce são manuscritos que foram organizados e publicados postumamente por projetos de universidades norte-americanas. Assim, CP é a abreviatura dos *Collected Papers* editados pela Universidade de Harvard, com o primeiro número correspondendo ao volume da coleção e o segundo ao(s) parágrafo(s) específico(s) da citação.