

MARCELO JOSÉ HANAUER

OS NOVOS REGIMES DA VERDADE

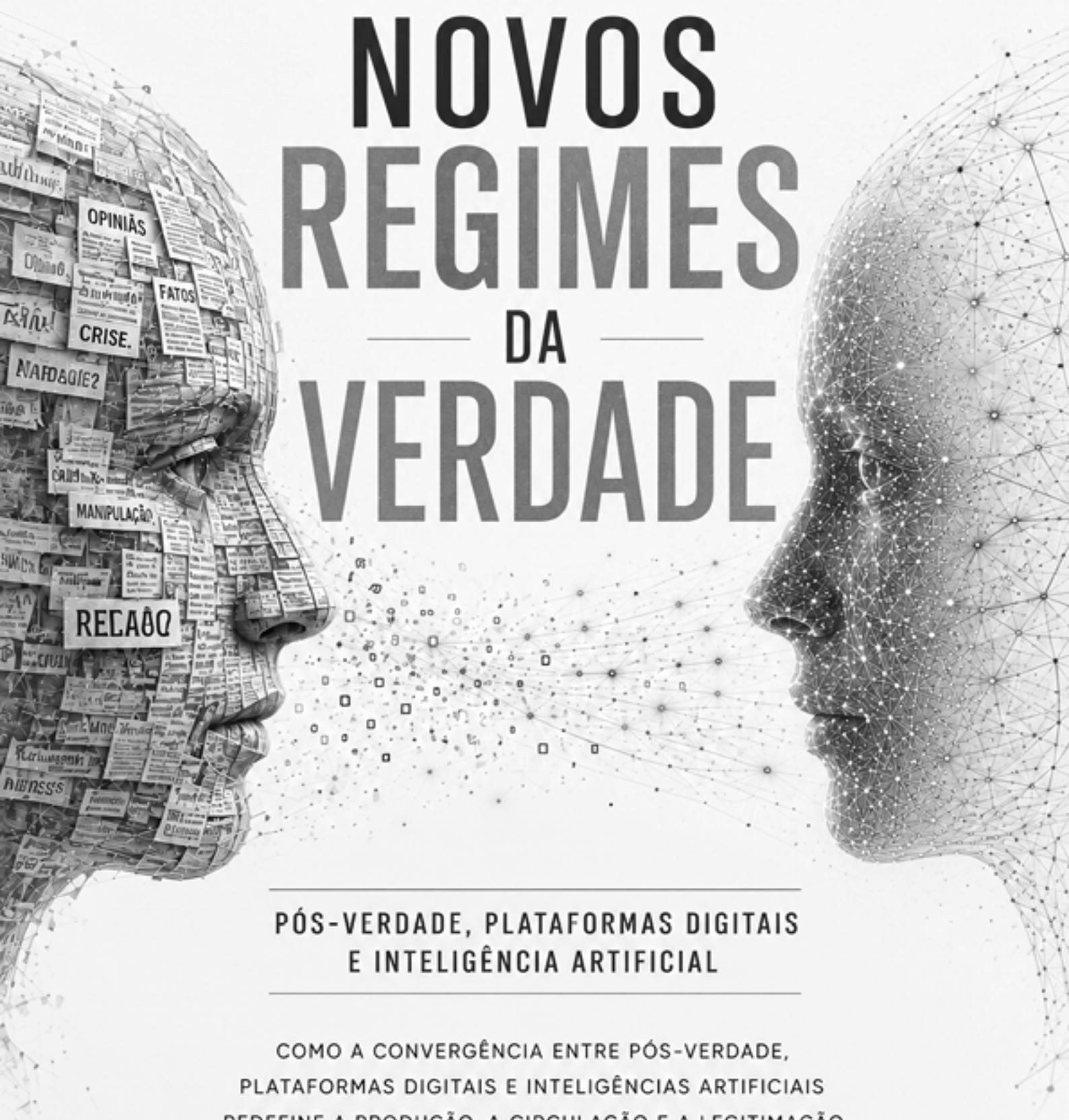


PÓS-VERDADE, PLATAFORMAS DIGITAIS
E INTELIGÊNCIA ARTIFICIAL

COMO A CONVERGÊNCIA ENTRE PÓS-VERDADE,
PLATAFORMAS DIGITAIS E INTELIGÊNCIAS ARTIFICIAIS
REDEFINE A PRODUÇÃO, A CIRCULAÇÃO E A LEGITIMAÇÃO
DO CONHECIMENTO NA SOCIEDADE CONTEMPORÂNEA

MARCELO JOSÉ HANAUER

OS NOVOS REGIMES DA VERDADE



PÓS-VERDADE, PLATAFORMAS DIGITAIS
E INTELIGÊNCIA ARTIFICIAL

COMO A CONVERGÊNCIA ENTRE PÓS-VERDADE,
PLATAFORMAS DIGITAIS E INTELIGÊNCIAS ARTIFICIAIS
REDEFINE A PRODUÇÃO, A CIRCULAÇÃO E A LEGITIMAÇÃO
DO CONHECIMENTO NA SOCIEDADE CONTEMPORÂNEA

1.ª Edição - Copyrights do texto - Autores e Autoras

Direitos de Edição Reservados à Editora Terried

É permitida a reprodução parcial ou total desta obra, desde que citada a fonte.



O conteúdo dos capítulos apresentados nesta obra são de inteira responsabilidade d@s autor@s, não representando necessariamente a opinião da Editora.

Permitimos a reprodução parcial ou total desta obra, considerado que seja citada a fonte e a autoria, além de respeitar a Licença Creative Commons indicada.

Conselho Editorial

Adilson Cristiano Habowski - ***Currículo Lattes***

Adilson Tadeu Basquerote Silva - ***Currículo Lattes***

Alexandre Carvalho de Andrade - ***Currículo Lattes***

Anísio Batista Pereira - ***Currículo Lattes***

Celso Gabatz - ***Currículo Lattes***

Cristiano Cunha Costa - ***Currículo Lattes***

Denise Santos Da Cruz - ***Currículo Lattes***

Emily Verônica Rosa da Silva Feijó - ***Currículo Lattes***

Fabiano Custódio de Oliveira - ***Currículo Lattes***

Fernanda Monteiro Barreto Camargo - ***Currículo Lattes***

Fredi dos Santos Bento - ***Currículo Lattes***

Guilherme Mendes Tomaz dos Santos - ***Currículo Lattes***

Humberto Costa - ***Currículo Lattes***

Leandro Antônio dos Santos - ***Currículo Lattes***

Lourenço Resende da Costa - ***Currículo Lattes***

Marcos Pereira dos Santos - ***Currículo Lattes***

Dados Internacionais de Catalogação na Publicação (CIP)
(Câmara Brasileira do Livro, SP, Brasil)

Os Novos Regimes da Verdade: Pós-verdade, Plataformas Digitais e Inteligência Artificial. Marcelo José Hanauer (Autor) -- Alegrete, RS : Editora Terried, 2026.

PDF
ISBN. 978-65-83367-95-2
1. Ensino

24-225451

CDD-918.17

Índices para catálogo sistemático:

1. Educação 90.14
2. Ensino 90.9



TERRIED

www.terried.com

contato@terried.com

SUMÁRIO

INTRODUÇÃO.....	6
CAPÍTULO 1.....	12
A VERDADE DEPOIS DOS FATOS: UMA INVESTIGAÇÃO SOBRE A CONSTITUIÇÃO DA PÓS-VERDADE	
CAPÍTULO 2.....	26
DAS ESTRUTURAS SOCIAIS ÀS PLATAFORMAS DIGITAIS: UMA INVESTIGAÇÃO SOBRE A EVOLUÇÃO DO CONCEITO DE REDES SOCIAIS	
CAPÍTULO 3.....	63
MÁQUINAS QUE PRODUZEM SIGNOS: UMA INVESTIGAÇÃO SOBRE A ORIGEM, O DESENVOLVIMENTO E OS EFEITOS CONTEMPORÂNEOS DAS INTELIGÊNCIAS ARTIFICIAIS GENERATIVAS	
CAPÍTULO 4.....	92
A CONVERGÊNCIA ENTRE PÓS-VERDADE, PLATAFORMAS DIGITAIS E INTELIGÊNCIAS ARTIFICIAIS GENERATIVAS	
CONCLUSÃO.....	112

INTRODUÇÃO

Nas últimas décadas, a humanidade passou a vivenciar uma transformação sem precedentes nas formas de produzir, compartilhar, acessar e validar o conhecimento. O desenvolvimento da internet, a consolidação das plataformas digitais e, mais recentemente, o avanço das inteligências artificiais generativas modificaram profundamente os processos de circulação da informação. Essas mudanças alteraram não apenas os meios pelos quais nos comunicamos, mas também as condições sociais, políticas, econômicas e culturais que sustentam a construção daquilo que reconhecemos como verdadeiro. Nesse cenário, a emergência da pós-verdade tornou-se um dos fenômenos mais debatidos nas Ciências Humanas e Sociais, justamente por evidenciar uma crise que ultrapassa a disseminação de informações falsas e alcança os próprios mecanismos de produção da confiança pública.

Foi essa inquietação que motivou a escrita desta obra. Muito se tem produzido sobre *fake news*, desinformação, algoritmos, redes sociais e inteligência artificial. Entretanto, grande parte dessa produção permanece fragmentada. Há estudos dedicados exclusivamente à pós-verdade, outros concentrados na arquitetura das plataformas digitais e um número crescente de pesquisas sobre inteligências artificiais generativas. Raramente, contudo, esses fenômenos são analisados como partes de uma mesma transformação histórica. Com frequência, as redes sociais são tratadas como causa exclusiva da desinformação; em outras ocasiões, a inteligência artificial aparece como uma ruptura tecnológica desvinculada das mudanças informacionais que a antecederam. Essa fragmentação dificulta a compreensão das relações entre esses processos e limita a construção de modelos explicativos capazes de interpretar, em sua complexidade, os desafios contemporâneos.

Este livro parte da hipótese de que os diferentes efeitos da pós-verdade não podem ser compreendidos isoladamente. Eles resultam da convergência entre mudanças epistemológicas, transformações sociotécnicas e novas formas de produção automatizada da informação. Pós-verdade, plataformas digitais e inteligências artificiais generativas integram um mesmo ecossistema informacional, no qual tecnologias, sujeitos, instituições, modelos econômicos e práticas cultu-

rais reorganizam continuamente as condições de produção, circulação, legitimação e contestação do conhecimento.

Essa perspectiva busca preencher uma lacuna na literatura contemporânea. Embora existam contribuições importantes sobre cada um desses temas, ainda são escassos os trabalhos que procuram articulá-los por meio de uma abordagem interdisciplinar capaz de integrar Filosofia, Sociologia, Ciência da Informação, Comunicação, Estudos de Plataforma, Inteligência Artificial e Educação. Mais do que descrever tecnologias ou revisar conceitos já consolidados, esta obra procura construir um modelo interpretativo que explique como a convergência entre plataformas digitais e inteligências artificiais generativas amplia, transforma e complexifica os efeitos da pós-verdade nas sociedades contemporâneas.

Em síntese, a obra defende que essa convergência configura um novo regime sociotécnico de produção da verdade. Nesse regime, algoritmos, plataformas e sistemas de inteligência artificial passam a integrar as infraestruturas responsáveis por mediar, legitimar e disputar aquilo que as sociedades contemporâneas reconhecem como verdadeiro. Essa proposição constitui a tese central da investigação e orienta o desenvolvimento dos capítulos seguintes.

Ao propor essa interpretação, o livro desloca o debate sobre a pós-verdade do campo exclusivamente discursivo para uma perspectiva sociotécnica. Compreender a crise contemporânea da verdade exige analisar conjuntamente sujeitos, instituições, plataformas digitais, algoritmos e inteligências artificiais como componentes de uma mesma arquitetura de produção do conhecimento. É dessa convergência que emergem os chamados novos regimes da verdade, eixo interpretativo que orienta toda a investigação.

Contribuições teóricas da obra

Esta obra não pretende apenas reunir conhecimentos produzidos sobre pós-verdade, plataformas digitais e inteligências artificiais generativas. Seu propósito é construir uma interpretação integradora que explique como esses fenômenos se articulam na reorganização contemporânea da produção, da circulação e da validação do conhecimento. Nesse sentido, a investigação apresenta cinco contribuições principais para o debate científico.

A primeira consiste em compreender a pós-verdade como um fenômeno simultaneamente epistemológico, sociotécnico e cultural. Em vez de restringi-la à disseminação de notícias falsas ou à manipulação política da informação, a obra a interpreta como expressão de uma transformação mais ampla nas condições pelas quais a verdade é socialmente produzida, reconhecida e contestada.

A segunda contribuição reside na reconstrução histórica do conceito de redes sociais. O livro demonstra que elas não surgiram com a internet nem podem ser confundidas com as plataformas digitais contemporâneas. Ao recuperar sua trajetória desde a Sociologia clássica até os atuais estudos sobre plataformização, busca-se conferir maior precisão conceitual a um termo frequentemente empregado de maneira indistinta nos debates público e acadêmico.

A terceira contribuição consiste em integrar a literatura sobre inteligências artificiais generativas ao campo dos estudos da pós-verdade. Em vez de compreender essas tecnologias apenas como ferramentas computacionais ou inovações isoladas, a investigação analisa sua participação na reorganização dos processos contemporâneos de produção simbólica, circulação da informação e construção da autoridade epistemológica.

A quarta contribuição refere-se à proposição de uma abordagem interdisciplinar para compreender a convergência entre pós-verdade, plataformas digitais e inteligências artificiais generativas. Esses fenômenos não devem ser analisados separadamente, pois constituem dimensões interdependentes de uma mesma transformação histórica, caracterizada pela crescente mediação algorítmica dos processos comunicacionais e cognitivos. A articulação entre Filosofia, Sociologia, Comunicação, Ciência da Informação, Estudos de Plataforma, Educação e Inteligência Artificial constitui um dos diferenciais analíticos desta obra.

Por fim, a quinta contribuição consiste na formulação de um modelo interpretativo para compreender a reorganização contemporânea dos regimes de produção da verdade. Plataformas digitais, algoritmos e inteligências artificiais generativas passaram a integrar as infraestruturas sociotécnicas responsáveis pela mediação do conhecimento, modificando profundamente as condições pe-

las quais as informações são produzidas, legitimadas e compartilhadas na sociedade contemporânea.

Essas contribuições não encerram o debate sobre os fenômenos analisados. Ao contrário, procuram oferecer um referencial teórico capaz de estimular novas investigações e favorecer o diálogo entre diferentes campos do conhecimento, reconhecendo que a compreensão da sociedade digital exige abordagens cada vez mais interdisciplinares e integradas.

Trata-se, portanto, de uma investigação teórica fundamentada em ampla revisão crítica da literatura nacional e internacional. Seu objetivo não é oferecer respostas definitivas para um fenômeno em permanente transformação, mas construir um percurso analítico que permita compreender suas múltiplas dimensões, identificar aproximações entre diferentes tradições de pesquisa e propor categorias conceituais que contribuam para o avanço do debate científico.

Para alcançar esse objetivo, o livro organiza-se em quatro capítulos articulados entre si.

O primeiro capítulo investiga a constituição histórica e epistemológica da pós-verdade, demonstrando que ela não corresponde simplesmente à existência de mentiras ou notícias falsas, mas à reorganização das condições pelas quais determinadas afirmações adquirem legitimidade social. São discutidas suas origens conceituais, seus fundamentos filosóficos e suas relações com a crise contemporânea de confiança nas instituições responsáveis pela produção do conhecimento.

O segundo capítulo reconstrói historicamente a evolução do conceito de redes sociais, desde suas bases na Sociologia clássica até sua transformação em plataformas digitais. Esse percurso evidencia que as redes sociais antecedem a internet em muitas décadas e que as plataformas atuais representam uma reorganização tecnológica de estruturas relacionais historicamente constituídas. Ao diferenciar rede social, mídia social, site de rede social e plataforma digital, busca-se oferecer maior precisão conceitual aos debates contemporâneos.

O terceiro capítulo dedica-se às inteligências artificiais generativas, examinando sua evolução histórica, seus fundamentos tecnológicos e seus impac-

tos sobre a produção contemporânea do conhecimento. São discutidas suas potencialidades para a pesquisa, a educação, a criatividade e a inovação, bem como os desafios éticos, epistemológicos e sociais decorrentes da produção automatizada de textos, imagens, áudios e vídeos.

O quarto capítulo, por sua vez, integra os três eixos anteriormente desenvolvidos e propõe uma interpretação interdisciplinar dos efeitos da pós-verdade impulsionados pelas plataformas digitais e pelas inteligências artificiais generativas. Defende-se que esses fenômenos não devem ser compreendidos como acontecimentos isolados, mas como expressões de uma nova ecologia informacional, caracterizada pela crescente mediação algorítmica dos processos de produção, circulação e validação do conhecimento.

Ao longo desse percurso, alguns conceitos serão progressivamente construídos e refinados. Entre eles, destacam-se a noção de pós-verdade como condição epistemológica e sociotécnica; a distinção entre redes sociais, mídias sociais, sites de redes sociais e plataformas digitais; a compreensão das inteligências artificiais generativas como participantes da reorganização da produção simbólica; e, sobretudo, a ideia de que vivemos uma profunda transformação dos regimes contemporâneos de produção da verdade. Esses conceitos não são apresentados como definições acabadas, mas como categorias analíticas desenvolvidas ao longo da investigação, cuja compreensão depende da articulação entre os diferentes capítulos.

Mais do que discutir tecnologias emergentes ou fenômenos comunicacionais específicos, esta obra propõe uma reflexão sobre uma questão fundamental para o século XXI: como as sociedades contemporâneas constroem aquilo que reconhecem como verdadeiro em um contexto no qual plataformas digitais, algoritmos e inteligências artificiais participam ativamente da produção e da circulação do conhecimento? Responder a essa pergunta exige ultrapassar explicações simplificadoras, reconhecer a complexidade dos fenômenos analisados e compreender que a crise contemporânea da verdade não decorre exclusivamente das tecnologias, mas da interação permanente entre sujeitos, instituições, infraestruturas digitais e novas formas de mediação algorítmica da informação.

Esta obra não pretende oferecer respostas definitivas para um fenômeno em permanente transformação. Busca, sobretudo, apresentar um referencial teórico capaz de ampliar a compreensão de uma questão que provavelmente acompanhará as próximas décadas da produção científica: de que maneira plataformas digitais e inteligências artificiais generativas estão reorganizando as condições pelas quais as sociedades contemporâneas constroem, legitimam e disputam aquilo que reconhecem como verdadeiro? É nesse horizonte que se inscreve esta investigação.

CAPÍTULO 1

A VERDADE DEPOIS DOS FATOS: UMA INVESTIGAÇÃO SOBRE A CONSTITUIÇÃO DA PÓS-VERDADE

A pós-verdade tornou-se uma das categorias mais recorrentes para interpretar as transformações contemporâneas da comunicação, da política e das relações sociais. Apesar de sua ampla circulação, o conceito não deve ser compreendido como a simples existência de mentiras, boatos ou informações incorretas. A falsificação dos fatos, a propaganda, a manipulação retórica e a fabricação de inimigos políticos atravessam toda a história humana. O que diferencia a pós-verdade é a formação de um ambiente social no qual a veracidade de uma afirmação deixa de constituir o principal critério para sua aceitação pública. Nesse ambiente, crenças, identidades, afetos, ressentimentos e vínculos de pertencimento podem exercer maior influência sobre a opinião das pessoas do que evidências verificáveis, demonstrações científicas ou registros documentais. A pós-verdade não designa, portanto, um período posterior à existência da verdade, mas uma condição cultural em que a verdade factual perde centralidade nas disputas pela interpretação da realidade (Araújo, 2020; Cesarino, 2021; McIntyre, 2018).

Essa distinção é fundamental porque impede que o conceito seja reduzido a um sinônimo sofisticado de mentira. A mentira tradicional reconhece, ainda que negativamente, a existência de uma verdade: alguém mente porque conhece ou presume conhecer determinado fato e procura ocultá-lo, alterá-lo ou substituí-lo deliberadamente. Na pós-verdade, por outro lado, o funcionamento discursivo pode dispensar essa relação estável com o verdadeiro. O enunciado passa a valer por sua capacidade de mobilizar uma comunidade, confirmar uma identidade, produzir indignação, fortalecer uma convicção anterior ou deslegitimar o adversário. Assim, uma alegação pode continuar sendo reproduzida mesmo depois de desmentida, pois sua eficácia não depende exclusivamente de sua correspondência com os acontecimentos. Ela permanece socialmente funcional enquanto oferece ao grupo uma narrativa emocionalmente satisfatória e politicamente conveniente (Genesini, 2018; Siebert; Pereira, 2020; Dominguez, 2023).

O termo *post-truth* apareceu em sua acepção contemporânea no ensaio “A Government of Lies”, publicado pelo dramaturgo sérvio-americano Steve Tesich na revista *The Nation*, em janeiro de 1992. Ao examinar a reação da sociedade estadunidense a acontecimentos como o escândalo de Watergate, a Guerra do Vietnã e conflitos posteriores, Tesich sugeriu que parte da população não era apenas enganada por seus governantes, mas começava a preferir narrativas confortáveis às verdades politicamente dolorosas. A novidade identificada pelo autor estava na disposição coletiva de aceitar a mentira como proteção contra fatos desagradáveis. A pós-verdade emergia, assim, não somente como estratégia de governantes, mas como uma espécie de pacto tácito entre produtores e consumidores de narrativas públicas: os primeiros ofereciam versões emocionalmente convenientes; os segundos aceitavam não submetê-las a exigências rigorosas de comprovação (Tesich, 1992).

Em 2004, Ralph Keyes ampliou esse diagnóstico com a publicação de *The Post-Truth Era: Dishonesty and Deception in Contemporary Life*. Keyes não restringiu a pós-verdade ao campo da propaganda política. Para ele, a relativização da honestidade disseminava-se também na vida cotidiana, na publicidade, nas relações profissionais, nos currículos, na comunicação empresarial e na produção da imagem pública dos indivíduos. O autor descreveu uma cultura na qual a mentira raramente se apresenta como mentira declarada. Em seu lugar, multiplicam-se eufemismos, omissões, meias-verdades, enquadramentos seletivos e justificativas morais que permitem ao sujeito enganar sem se reconhecer plenamente como mentiroso. Essa normalização reduz o constrangimento social associado ao engano e torna mais fluida a fronteira entre informação, interpretação interessada e fabricação discursiva (Keyes, 2004).

A consolidação internacional do conceito ocorreu em 2016, quando a Oxford Dictionaries escolheu *post-truth* como palavra do ano. A instituição definiu o termo como relativo a circunstâncias nas quais fatos objetivos possuem menos influência na formação da opinião pública do que apelos à emoção e às crenças pessoais. A escolha esteve relacionada, sobretudo, à intensificação do uso da expressão durante o referendo que decidiu a saída do Reino Unido da União Europeia, conhecido como Brexit, e durante a eleição presidencial dos Estados Unidos. A relevância daquele momento não significa que a pós-verdade tenha

surgido em 2016, mas que condições políticas, comunicacionais e tecnológicas tornaram o fenômeno particularmente visível e permitiram sua entrada definitiva no vocabulário público e acadêmico internacional (Oxford Languages, 2016; McIntyre, 2018).

O prefixo “pós”, nesse contexto, não deve ser interpretado exclusivamente em sentido cronológico. Pós-verdade não significa que a humanidade tenha superado uma suposta era anterior de plena transparência, objetividade ou consenso factual. Sociedades passadas também produziram propaganda, censura, perseguições, mitologias políticas e falsificações. O “pós” aponta, antes, para a perda de relevância normativa da verdade em determinados espaços da vida pública. Não se afirma que os fatos deixaram de existir, mas que sua existência já não garante sua capacidade de organizar o debate coletivo. O problema central desloca-se da oposição simples entre verdade e falsidade para a disputa sobre quais instituições, pessoas e procedimentos possuem legitimidade para determinar o que deve ser reconhecido como verdadeiro (Oliveira; Melo, 2024; Araújo, 2020).

Essa crise de legitimidade atinge aquilo que Letícia Cesarino denomina “sistema de peritos”: o conjunto formado por universidades, imprensa profissional, instituições científicas, agências públicas, especialistas, organismos internacionais e procedimentos técnicos responsáveis pela produção e validação de conhecimentos. Durante boa parte da modernidade, essas instituições desempenharam uma função de mediação entre fenômenos complexos e a compreensão pública. Contudo, a intensificação da digitalização, a fragmentação das audiências e a personalização algorítmica da informação contribuíram para enfraquecer a autoridade desses mediadores. Na condição de pós-verdade, o especialista deixa de ser reconhecido automaticamente por sua formação ou posição institucional e passa a concorrer, em relativa igualdade comunicacional, com influenciadores, celebridades, lideranças religiosas, militantes, produtores anônimos e conteúdos não verificados (Cesarino, 2021).

Esse processo não deve ser interpretado como se toda crítica às instituições especializadas fosse irracional. A ciência, o jornalismo, o Estado e as universidades não são estruturas neutras, infalíveis ou imunes a interesses econômicos, políticos, coloniais e sociais. A história registra exclusões, erros, silenciamentos e conhecimentos produzidos a partir de perspectivas dominantes. Por essa razão,

o enfrentamento da pós-verdade não pode se transformar em defesa autoritária de uma ciência que simplesmente exige obediência do público. A confiança social depende de transparência metodológica, reconhecimento das incertezas, abertura ao contraditório e capacidade de incluir diferentes sujeitos na produção do conhecimento. Caso contrário, a própria atitude autoritária das instituições pode intensificar a desconfiança que pretende combater (Rosa, 2020; Lima et al., 2020).

A pós-verdade desenvolve-se, portanto, em uma tensão. De um lado, há uma crítica legítima aos limites das instituições modernas e aos modos históricos pelos quais determinadas vozes foram desautorizadas. De outro, agentes políticos e econômicos podem instrumentalizar essa crítica para desacreditar qualquer conhecimento inconveniente aos seus interesses. A passagem da crítica epistemológica para o negacionismo ocorre quando a constatação de que todo conhecimento é produzido em condições históricas é indevidamente convertida na afirmação de que todas as versões possuem o mesmo valor. Reconhecer que a ciência é uma prática social não significa negar sua capacidade de produzir conhecimentos mais confiáveis por meio de observação, argumentação pública, revisão por pares, replicação, confronto de hipóteses e correção sistemática de erros (Mcintyre, 2018; Araújo, 2020).

A pós-verdade também não se confunde integralmente com as chamadas *fake news*. Notícias falsas são conteúdos que imitam características do jornalismo ou da informação documental para transmitir afirmações enganosas, descontextualizadas ou fabricadas. A pós-verdade é uma condição mais ampla, na qual esses conteúdos podem adquirir força e circular com eficácia. Em outras palavras, a notícia falsa constitui um objeto comunicacional; a pós-verdade constitui um ambiente epistemológico, afetivo, político e tecnológico que favorece sua aceitação. Uma informação falsa isolada não cria necessariamente uma condição de pós-verdade. Para isso, é necessário que existam comunidades dispostas a aceitá-la, redes capazes de distribuí-la, lideranças que a legitimem e estruturas de crença que a tornem plausível para determinado público (Kegler; Santos, 2020; Araújo, 2020).

Do mesmo modo, é importante diferenciar informação incorreta de desinformação. A informação incorreta pode ser compartilhada por engano, desconhecimento ou interpretação equivocada. A desinformação envolve a produção ou a

circulação estratégica de conteúdos falsos, distorcidos ou manipulados com a finalidade de enganar, influenciar comportamentos ou causar danos. A pós-verdade pode envolver ambas, pois conteúdos inicialmente produzidos de modo intencional podem ser retransmitidos por sujeitos que acreditam sinceramente neles. Essa passagem entre fabricação deliberada e adesão espontânea é um dos elementos que tornam o fenômeno complexo. A mentira estratégica deixa de depender apenas de seus criadores quando passa a ser defendida afetivamente por comunidades que a incorporam à própria identidade (Bertoli; Ribeiro, 2025; Giordani et al., 2021).

As redes sociais possuem papel decisivo nesse desenvolvimento, embora não possam ser consideradas a causa única da pós-verdade. Elas reorganizam a produção e a circulação da informação ao eliminar parte das barreiras tradicionais de publicação, acelerar o compartilhamento e permitir que qualquer usuário participe simultaneamente como receptor, produtor e distribuidor de conteúdos. Essa horizontalização possui um aspecto democrático, pois amplia a possibilidade de expressão pública. Contudo, também reduz a capacidade dos antigos mecanismos editoriais de verificar previamente a informação. No mesmo espaço de visibilidade podem circular reportagens investigativas, opiniões pessoais, sátiras, propaganda, relatos legítimos, teorias conspiratórias e falsificações deliberadas, frequentemente apresentados em formatos semelhantes e consumidos sem distinção clara de origem ou qualidade (Farias, 2020; Kegler; Santos, 2020).

Além disso, as plataformas digitais não organizam os conteúdos de forma neutra ou apenas cronológica. Sistemas algorítmicos selecionam aquilo que será mostrado com base em dados de comportamento, preferências presumidas, histórico de interação e probabilidade de engajamento. Esse modelo econômico transforma atenção, permanência e interação em recursos comercializáveis. Conteúdos capazes de provocar surpresa, medo, raiva ou indignação tendem a gerar respostas intensas, sendo potencialmente favorecidos em circuitos orientados pelo engajamento. A infraestrutura técnica das plataformas não inventa os conflitos sociais, mas pode ampliá-los ao recompensar conteúdos emocionalmente mobilizadores, polarizados e facilmente compartilháveis (Cesarino, 2021; Farias, 2020).

Nesse ambiente, a autoridade de uma informação passa a ser frequentemente confundida com sua popularidade. Número de visualizações, compartilhamentos, comentários e seguidores pode ser interpretado como evidência de verdade, mesmo quando representa apenas alcance ou intensidade de circulação. A repetição produz familiaridade, e a familiaridade pode ser convertida psicologicamente em plausibilidade. Uma afirmação vista muitas vezes pode parecer intuitivamente verdadeira, ainda que sua fonte original seja desconhecida ou tenha sido desmentida. A força da pós-verdade reside, em parte, nessa transformação da recorrência em aparência de confirmação: o conteúdo circula porque parece verdadeiro e parece verdadeiro porque circula amplamente (Mcintyre, 2018; Dominguez, 2023).

A dimensão afetiva é igualmente indispensável para compreender o fenômeno. Os sujeitos não recebem informações como observadores neutros. Eles as interpretam segundo experiências anteriores, valores, medos, expectativas, vínculos sociais e posições políticas. As pessoas tendem a aceitar com menor resistência informações compatíveis com suas crenças e a avaliar com maior rigor aquelas que ameaçam suas identidades. Esse funcionamento não significa que todos sejam incapazes de raciocinar, mas que a razão opera em relação constante com afetos e pertencimentos. Na pós-verdade, atores estrategicamente interessados exploram essas disposições, oferecendo narrativas que transformam problemas complexos em explicações simples e emocionalmente coerentes (Mcintyre, 2018; Miranda; Caldas, 2021).

Essa dinâmica ajuda a compreender por que a simples apresentação de dados corretos nem sempre é suficiente para modificar uma crença falsa. Quando determinada convicção integra a identidade social do indivíduo, abandoná-la pode significar perder pertencimento, admitir ter sido enganado ou aproximar-se simbolicamente de um grupo percebido como adversário. A disputa deixa de ocorrer apenas no plano da informação e passa a envolver reconhecimento, lealdade e proteção da autoimagem. Consequentemente, uma correção factual pode ser recebida como agressão política ou moral, sobretudo quando é apresentada de maneira humilhante, autoritária ou descontextualizada (Cesarino, 2021; Miranda; Caldas, 2021).

A pós-verdade constitui também um paradigma argumentativo. Em uma concepção clássica do debate público, os interlocutores deveriam apresentar ra-

zões, evidências e justificativas capazes de persuadir inclusive aqueles que não compartilham suas posições iniciais. Na lógica da pós-verdade, entretanto, grande parte da comunicação não busca convencer o adversário, mas reafirmar os já convencidos. O discurso dirige-se ao próprio grupo, fortalece fronteiras identitárias e produz sinais de fidelidade. A argumentação converte-se em performance de pertencimento. Seu êxito não é medido pela capacidade de construir entendimento comum, mas pela intensidade com que mobiliza apoiadores e desqualifica opositores (Dominguez, 2023).

Por essa razão, a pós-verdade é particularmente funcional em contextos de polarização política. A sociedade passa a ser representada como conflito entre comunidades moralmente incompatíveis, nas quais cada grupo atribui ao outro não apenas erro de interpretação, mas desonestidade, ignorância ou perversidade. Uma vez destruído o mínimo reconhecimento mútuo, as fontes aceitas por um campo tornam-se automaticamente suspeitas para o outro. A possibilidade de estabelecer fatos comuns diminui, pois as evidências passam a ser avaliadas prioritariamente a partir da identidade de quem as apresenta. O resultado é a formação de realidades informacionais paralelas, sustentadas por fontes, autoridades e repertórios narrativos distintos (Farias, 2020; Bertoli; Ribeiro, 2025).

Os usos políticos da pós-verdade incluem a fabricação de inimigos, o ataque à imprensa, a contestação preventiva de resultados eleitorais, a disseminação de teorias conspiratórias e a deslegitimação de instituições de controle. Um agente político não precisa demonstrar integralmente sua versão; pode produzir dúvida suficiente para impedir a consolidação de um entendimento comum. Essa estratégia opera pela saturação do espaço público com explicações concorrentes, acusações recíprocas e controvérsias intermináveis. Quando todas as fontes parecem suspeitas e nenhuma versão parece verificável, a população pode abandonar a busca por critérios objetivos e aderir à narrativa do líder ou do grupo com o qual mantém maior identificação (Bertoli; Ribeiro, 2025; McIntyre, 2018).

Um dos efeitos mais profundos desse funcionamento é a produção de cinismo. A finalidade da desinformação nem sempre consiste em fazer o público acreditar plenamente em uma mentira específica. Em determinados casos, basta convencê-lo de que todos mentem, de que nenhuma instituição merece confiança e de que a verdade é inalcançável. Esse relativismo generalizado favorece os

agentes com maior poder de comunicação, pois elimina critérios externos pelos quais suas afirmações poderiam ser avaliadas. Quando a verdade pública é reduzida a uma disputa entre versões equivalentes, a capacidade de exigir responsabilidade dos governantes, das empresas e das instituições torna-se enfraquecida (Araújo, 2020; Cesarino, 2021).

As consequências tornam-se especialmente graves quando a pós-verdade alcança áreas como saúde, ciência e políticas públicas. Durante emergências sanitárias, informações falsas podem orientar decisões sobre vacinação, tratamentos, prevenção e confiança nos profissionais. Nessas situações, narrativas científicas disputam atenção com depoimentos pessoais, soluções milagrosas, interpretações conspiratórias e conteúdos pseudocientíficos. A experiência da pandemia de covid-19 evidenciou que a infodemia — o excesso de informações corretas, imprecisas e falsas circulando simultaneamente — dificulta a identificação de orientações confiáveis e compromete a resposta coletiva aos riscos (Giordani et al., 2021).

No campo da ciência, a pós-verdade não se limita à rejeição de resultados específicos. Ela pode comprometer a confiança nos próprios procedimentos pelos quais o conhecimento é produzido. Questionar um estudo, examinar métodos e exigir transparência são práticas indispensáveis à ciência. O negacionismo, contudo, utiliza a existência normal de incertezas e divergências científicas para sugerir que nenhum conhecimento é suficientemente confiável. A dúvida, que na investigação científica é uma ferramenta metodológica, converte-se em estratégia retórica de paralisia. Exige-se da ciência uma certeza absoluta que nenhuma atividade humana pode oferecer e, diante da ausência dessa certeza, justifica-se a adesão a crenças desprovidas de sustentação equivalente (Lima et al., 2020; Cruz Junior, 2019).

Os usos comerciais da pós-verdade também merecem atenção. Publicidade disfarçada de informação, depoimentos fabricados, métricas artificiais de popularidade, influenciadores que ocultam relações comerciais e empresas que produzem dúvida sobre evidências científicas demonstram que a manipulação da verdade não constitui apenas instrumento eleitoral. A economia da atenção cria incentivos para a publicação rápida de conteúdos atraentes, enquanto a correção posterior costuma receber menos visibilidade. Dessa forma, os custos

sociais da desinformação são distribuídos coletivamente, mas seus benefícios econômicos podem ser apropriados por plataformas, anunciantes, produtores de conteúdo e agentes especializados em monetizar engajamento (Keyes, 2004; Kegler; Santos, 2020).

A emergência das inteligências artificiais generativas acrescenta uma nova camada a esse problema. Esses sistemas podem produzir textos, imagens, áudios e vídeos convincentes em escala, com baixo custo e elevada velocidade. Embora ofereçam possibilidades importantes para educação, pesquisa, acessibilidade e criação, também podem ser empregados para automatizar propaganda, fabricar testemunhos, simular autoridades, personalizar mensagens enganosas e inundar o ambiente informacional com conteúdos sintéticos. A questão central não consiste em afirmar que a inteligência artificial cria sozinha a pós-verdade, mas reconhecer que ela amplia a capacidade humana de produzir versões plausíveis da realidade sem que exista necessariamente um acontecimento correspondente.

A inteligência artificial generativa modifica ainda o estatuto da prova audiovisual. Fotografias, gravações e vídeos foram historicamente tratados, mesmo com ressalvas, como registros dotados de força documental. Com a disseminação de conteúdos sintéticos realistas, não apenas imagens falsas podem ser aceitas como verdadeiras, mas imagens autênticas podem ser rejeitadas sob a alegação de terem sido artificialmente produzidas. Forma-se, assim, um duplo risco epistemológico: a fabricação convincente do que nunca aconteceu e a negação estratégica do que efetivamente ocorreu. Esse segundo movimento é particularmente relevante, pois permite que agentes confrontados com evidências reais as classifiquem como montagens, aprofundando a erosão da confiança pública.

Nesse sentido, a pós-verdade não significa apenas acreditar no falso. Ela pode significar a incapacidade crescente de estabelecer socialmente quais procedimentos permitem distinguir afirmações mais ou menos confiáveis. A crise não se encontra exclusivamente no conteúdo das mensagens, mas na infraestrutura de validação do conhecimento. Quando autoria, fonte, contexto, método e cadeia de produção tornam-se opacos, o receptor precisa decidir em condições de incerteza. Em tais situações, identidade, emoção e familiaridade assumem o lugar dos critérios de verificação. A pós-verdade constitui, portanto, uma crise dos dispositivos coletivos de confiança e não apenas uma falha individual de julgamento (Araújo, 2020; Cesarino, 2021).

Enfrentar esse fenômeno exige superar respostas excessivamente simples. A checagem de fatos é necessária, mas insuficiente quando atua somente depois que conteúdos falsos já alcançaram milhões de pessoas ou foram incorporados a identidades coletivas. A remoção de conteúdos pode ser necessária em casos específicos, mas também envolve debates legítimos sobre liberdade de expressão, transparência, poder privado das plataformas e possibilidade de abuso. A educação midiática é indispensável, porém não pode transferir toda a responsabilidade ao usuário, como se indivíduos isolados fossem capazes de enfrentar estruturas algorítmicas, campanhas organizadas e modelos econômicos desenhados para capturar sua atenção (Luiz; Almeida, 2023; Gomes, 2021).

Uma resposta consistente requer combinação de medidas institucionais, tecnológicas, educacionais e culturais. Isso inclui transparência sobre critérios algorítmicos, identificação de conteúdos sintéticos, responsabilização proporcional de agentes que lucram com campanhas de desinformação, fortalecimento do jornalismo profissional, apoio à divulgação científica, formação para leitura crítica das mídias e criação de espaços de diálogo capazes de reconstruir algum terreno factual comum. Também exige que instituições científicas e públicas comuniquem incertezas, métodos e limites de forma compreensível, sem abandonar o rigor e sem tratar o público como cognitivamente incapaz (Lima et al., 2020; Cesarino, 2021).

A educação para o enfrentamento da pós-verdade não deve restringir-se ao ensino de técnicas para identificar notícias falsas. É necessário compreender como se constroem evidências, como operam as fontes, quais interesses atravessam a circulação da informação e por que determinados conteúdos despertam adesão emocional. O estudante precisa aprender a perguntar quem produziu uma afirmação, com quais métodos, em que contexto, para qual público e com quais possibilidades de contestação. Precisa também reconhecer seus próprios vínculos afetivos com as informações, pois a desinformação mais eficaz não é necessariamente aquela que parece absurda, mas aquela que confirma exatamente aquilo que o sujeito deseja ou teme que seja verdadeiro.

Uma abordagem inovadora deve considerar, ainda, que a pós-verdade não representa somente degradação da racionalidade, mas transformação da forma de constituir pertencimentos. Narrativas falsas ou distorcidas podem funcionar

como instrumentos de reconhecimento coletivo: ao compartilhá-las, os sujeitos demonstram fidelidade, marcam posição e confirmam sua inserção em determinado grupo. Por isso, desmentir uma informação sem compreender a função social que ela exerce pode ser pouco eficaz. O enfrentamento precisa oferecer não apenas fatos, mas outras possibilidades de participação, reconhecimento, identidade e elaboração dos conflitos sociais.

Nessa perspectiva, a verdade pública não deve ser concebida como imposição vertical de uma autoridade infalível, tampouco como simples soma de opiniões individuais. Ela é resultado provisório de procedimentos coletivos de verificação, crítica, documentação, debate e correção. Sua força não decorre da ausência de conflitos, mas da existência de regras pelas quais as afirmações podem ser examinadas publicamente. Defender a verdade em uma sociedade democrática não significa eliminar a pluralidade de interpretações; significa preservar a possibilidade de distinguir interpretação, erro, incerteza, opinião e fabricação deliberada.

Conclui-se que a pós-verdade constitui uma condição histórica em que a disputa pela realidade é reorganizada pela articulação entre emoções, identidades, crise institucional, economia da atenção, plataformas digitais e tecnologias de produção automatizada de conteúdo. Sua origem conceitual antecede as redes sociais, mas seu desenvolvimento contemporâneo é inseparável delas. A pós-verdade não elimina os fatos, porém enfraquece sua capacidade de produzir consequências comuns. Não extingue a verdade, mas fragmenta os procedimentos sociais que permitiam reconhecê-la. O desafio científico, educacional e democrático consiste, assim, em reconstruir condições compartilhadas de confiança sem restaurar modelos autoritários de autoridade e sem ceder ao relativismo que transforma toda afirmação em mera versão. Essa reconstrução exige rigor factual, abertura crítica, responsabilidade tecnológica e compreensão profunda dos vínculos afetivos e sociais que sustentam a circulação do verdadeiro e do falso.

REFERÊNCIAS

ARAÚJO, Carlos Alberto Ávila. O fenômeno da pós-verdade e suas implicações para a agenda de pesquisa na Ciência da Informação. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 25, p. 1-17, 2020. DOI: <https://doi.org/10.5007/1518-2924.2020.e72673>. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2020.e72673>. Acesso em: 14 jul. 2026.

BERTOLI, Juliana Martins; RIBEIRO, Manuella Maia. Desinformação, pós-verdade e instabilidade democrática. **Organizações & Sociedade**, Salvador, v. 32, 2025. Disponível em: <https://www.scielo.br/j/osoc/a/4Jjx58YTgTcFyc99SmFx-j6C/>. Acesso em: 14 jul. 2026.

CESARINO, Letícia. Pós-verdade e a crise do sistema de peritos: uma explicação cibernética. **Ilha: Revista de Antropologia**, Florianópolis, v. 23, n. 1, p. 73-96, 2021. DOI: <https://doi.org/10.5007/2175-8034.2021.e75630>. Disponível em: <https://periodicos.ufsc.br/index.php/ilha/article/view/75630>. Acesso em: 14 jul. 2026.

CRUZ JUNIOR, Gilson. Pós-verdade: a nova guerra contra os fatos em tempos de fake news. **ETD – Educação Temática Digital**, Campinas, v. 21, n. 1, p. 278-284, 2019. Disponível em: https://educa.fcc.org.br/scielo.php?pid=S1676-25922019000100278&script=sci_arttext. Acesso em: 14 jul. 2026.

DOMINGUEZ, Maria Garcia de Almeida. A pós-verdade como paradigma argumentativo. **Linha D'Água**, São Paulo, v. 36, n. 2, 2023. Disponível em: <https://revistas.usp.br/linhadagua/article/view/204983>. Acesso em: 14 jul. 2026.

FARIAS, Luiz Alberto de. Comunicação, opinião pública e os impactos da revolução digital na era da pós-verdade e fake news. **Organicom**, São Paulo, v. 17, n. 34, p. 71-81, 2020. Disponível em: <https://revistas.usp.br/organicom/article/view/176133>. Acesso em: 14 jul. 2026.

GENESINI, Silvio. A pós-verdade é uma notícia falsa. **Revista USP**, São Paulo, n. 116, p. 45-58, 2018. DOI: <https://doi.org/10.11606/issn.2316-9036.v0i116p45-58>. Disponível em: <https://revistas.usp.br/revusp/article/view/146577>. Acesso em: 14 jul. 2026.

GIORDANI, Rubia Carla Formighieri; DONASOLO, João Pedro Gambetta; AMES, Vanessa Dias; GIORDANI, Rosane Cássia de Freitas. A ciência entre a infodemia e outras narrativas da pós-verdade: desafios em tempos de pandemia. **Ciência & Saúde Coletiva**, Rio de Janeiro, v. 26, n. 7, p. 2863-2872, 2021. DOI: <https://doi.org/10.1590/1413-81232021267.05892021>. Disponível em: <https://www.scielo.br/j/csc/a/MWfcvZ797BYyNSJBQTpNP8K/>. Acesso em: 14 jul. 2026.

GOMES, Camila Paula de Barros. O impacto das fake news sobre as políticas públicas. **Revista Digital de Direito Administrativo**, Ribeirão Preto, v. 8, n. 2, p. 191-206, 2021. Disponível em: <https://revistas.usp.br/rdda/article/view/179180>. Acesso em: 14 jul. 2026.

KEGLER, Bruno; SANTOS, Nina. Sociedade de plataformas, pós-verdade e fake news: um debate sobre a opinião pública possível. **Organicom**, São Paulo, v. 17, n. 34, 2020. Disponível em: <https://revistas.usp.br/organicom/article/download/170515/168690>. Acesso em: 14 jul. 2026.

KEYES, Ralph. *The post-truth era: dishonesty and deception in contemporary life*. New York: St. Martin's Press, 2004.

LIMA, Nathan Willig; VAZATA, Pedro Antonio Viana; MORAES, Andrei Steinbach; OSTERMANN, Fernanda; CAVALCANTI, Cláudio José de Holanda. Educação em Ciências nos tempos de pós-verdade: reflexões metafísicas a partir dos estudos das Ciências de Bruno Latour. **Revista Brasileira de Pesquisa em Educação em Ciências**, Belo Horizonte, v. 19, p. 155-189, 2019. DOI: <https://doi.org/10.28976/1984-2686rbpec2019u155189>.

LUIZ, Thiago Cury; ALMEIDA, Marco Antônio de. Quando a pós-verdade e o populismo se encontram: intercorrências de pós-verdade, populismo e fact-checking. **Galáxia**, São Paulo, n. 48, 2023. Disponível em: <https://www.scielo.br/j/gal/a/yZHW8tDBTTSVtNggXZ8VZJm/>. Acesso em: 14 jul. 2026.

MCINTYRE, Lee. **Post-truth**. Cambridge, Massachusetts: The MIT Press, 2018. Disponível em: <https://mitpress.mit.edu/9780262535045/post-truth/>. Acesso em: 14 jul. 2026.

MIRANDA, Luciana Lobo; CALDAS, Heloisa. Considerações psicanalíticas sobre a pós-verdade e as fake news. **Revista Latinoamericana de Psicopatologia Fundamental**, São Paulo, v. 24, 2021. Disponível em: <https://www.scielo.br/j/rlpf/a/xgJscBpHgZPcyRmxqDrqpDD/>. Acesso em: 14 jul. 2026.

OLIVEIRA, Amurabi; MELO, Marina Félix de. As ciências sociais e os cientistas no contexto de pós-verdade. **Sociedade e Estado**, Brasília, 2024. Disponível em: <https://www.scielo.br/j/se/a/V5LzpCsZKQmdpyZ7hG3h8Cv/>. Acesso em: 14 jul. 2026.

OXFORD LANGUAGES. **Word of the year 2016: post-truth**. Oxford: Oxford University Press, 2016. Disponível em: <https://global.oup.com/academic/content/word-of-the-year/>. Acesso em: 14 jul. 2026.

ROSA, Katemari. Pós-verdade para quem? Fatos produzidos por uma ciência racista. **Caderno Brasileiro de Ensino de Física**, Florianópolis, v. 37, n. 3, p. 1440-1468, 2020. DOI: <https://doi.org/10.5007/2175-7941.2020v37n3p1440>. Disponível em: <https://periodicos.ufsc.br/index.php/fisica/article/view/74989>. Acesso em: 14 jul. 2026.

SIEBERT, Silvânia; PEREIRA, Israel Vieira. A pós-verdade como acontecimento discursivo. **Linguagem em (Dis)curso**, Tubarão, v. 20, n. 2, p. 239-249, 2020. DOI: <https://doi.org/10.1590/1982-4017-200201-00-00>. Disponível em: <https://www.scielo.br/j/ld/a/vykt83t8h8874gJT7ys46sy/>. Acesso em: 14 jul. 2026.

TESICH, Steve. *A government of lies*. **The Nation**, New York, v. 254, n. 1, p. 12–14, 6 jan. 1992. Disponível em: <https://www.thefreelibrary.com/A+government+of+lies.-a011665982>. Acesso em: 14 jul. 2026.

CAPÍTULO 2

DAS ESTRUTURAS SOCIAIS ÀS PLATAFORMAS DIGITAIS: UMA INVESTIGAÇÃO SOBRE A EVOLUÇÃO DO CONCEITO DE REDES SOCIAIS

Quando se utiliza a expressão “redes sociais” no debate contemporâneo, é comum que ela seja imediatamente associada a plataformas digitais como *Facebook*, *Instagram*, *X* (antigo *Twitter*), *TikTok* ou *LinkedIn*. Essa associação, embora compreensível diante da centralidade que tais ambientes adquiriram na vida cotidiana, produz uma redução conceitual significativa. A literatura científica demonstra que as redes sociais não surgiram com a internet, tampouco constituem uma invenção das empresas de tecnologia do século XXI. Antes de se converterem em infraestruturas digitais, as redes sociais já eram objeto de investigação da Sociologia, da Antropologia, da Psicologia Social e, posteriormente, da Ciência da Informação e da Comunicação. O desenvolvimento das plataformas digitais representa apenas um estágio recente de um conceito cuja origem está vinculada ao estudo das relações humanas, das formas de organização da vida coletiva e dos mecanismos pelos quais indivíduos, grupos e instituições estabelecem conexões, produzem cooperação, exercem influência e compartilham recursos materiais e simbólicos (Barnes, 1954; Wellman, 1988; Recuero, 2018).

Essa constatação exige uma distinção metodológica fundamental. As redes sociais não devem ser confundidas com as tecnologias que, em diferentes momentos históricos, serviram de suporte para sua manifestação. Relações familiares, comunidades religiosas, corporações profissionais, vizinhanças, associações políticas, instituições educacionais e redes comerciais constituem exemplos históricos de redes sociais muito anteriores ao advento da internet. O que caracteriza uma rede não é o meio pelo qual ela se comunica, mas a existência de atores interligados por vínculos relativamente estáveis, capazes de produzir circulação de informações, bens, afetos, normas, confiança e poder. As tecnologias de comunicação modificam profundamente a escala, a velocidade e a intensidade dessas conexões, porém não criam, por si mesmas, o fenômeno social da rede (Castells, 1999; Wasserman; Faust, 1994).

A compreensão contemporânea das redes sociais resulta de um longo percurso intelectual que atravessa diferentes tradições das ciências humanas. Ainda no final do século XIX, Georg Simmel (2006) introduziu uma das contribuições mais importantes para esse campo ao deslocar o olhar sociológico das grandes instituições para as formas de interação entre indivíduos. Em vez de compreender a sociedade como uma entidade abstrata ou exclusivamente determinada por estruturas políticas e econômicas, Simmel propôs que ela fosse analisada como um processo permanente de associação. A sociedade não existiria acima dos indivíduos, mas seria continuamente produzida pelas múltiplas relações que eles estabelecem entre si. Essa perspectiva inaugura uma mudança epistemológica decisiva: o objeto central da Sociologia deixa de ser apenas o grupo ou a instituição e passa a incluir a própria configuração das relações sociais.

Ao enfatizar que a sociedade é produzida pelas interações, Simmel rompe com concepções excessivamente substancialistas da vida social. Seu interesse dirige-se menos aos indivíduos isoladamente considerados e mais às formas pelas quais eles se vinculam. Cooperação, conflito, competição, dominação, subordinação, troca e sociabilidade tornam-se categorias fundamentais para compreender a dinâmica social. Embora o autor jamais tenha utilizado a expressão “rede social” na acepção contemporânea, suas reflexões forneceram o fundamento teórico que permitiria, décadas depois, o surgimento da Análise de Redes Sociais (*Social Network Analysis*). A ideia de que a posição ocupada por um indivíduo nas relações pode ser tão importante quanto seus atributos pessoais permanece como um dos princípios organizadores desse campo de estudos (Simmel, 2006; Wellman, 1988).

As contribuições de Simmel ganharam novos desdobramentos ao longo da primeira metade do século XX, especialmente com o trabalho do psiquiatra e sociólogo Jacob Levy Moreno. Interessado em compreender os vínculos afetivos estabelecidos em grupos humanos, Moreno (1934) desenvolveu o método sociométrico, cuja finalidade consistia em representar graficamente as relações existentes entre os membros de uma coletividade. Em vez de limitar-se à descrição das características individuais, o pesquisador buscava identificar quem escolhia quem, quem permanecia isolado, quem ocupava posições centrais e quais grupos se formavam espontaneamente dentro de organizações sociais. Nasciam,

assim, os primeiros sociogramas, representações visuais que permitiam observar padrões de relacionamento anteriormente invisíveis.

A inovação metodológica introduzida por Moreno ultrapassou o simples uso de diagramas. Pela primeira vez, as relações sociais passaram a ser tratadas como estruturas passíveis de mensuração, comparação e análise sistemática. O foco deslocava-se definitivamente das propriedades individuais para os padrões de conexão existentes entre os sujeitos. Conceitos hoje amplamente difundidos — como centralidade, isolamento, reciprocidade e coesão — possuem raízes nesse esforço pioneiro de representar matematicamente as relações humanas. Ainda que as técnicas contemporâneas de análise de redes sejam muito mais sofisticadas, o princípio fundamental permanece o mesmo: compreender a organização da vida social exige examinar a arquitetura das conexões e não apenas os atributos dos indivíduos que delas participam (Moreno, 1934; Wasserman; Faust, 1994).

Nas décadas seguintes, a Antropologia Social britânica ampliou significativamente esse horizonte analítico. Investigando comunidades locais, processos migratórios e relações de parentesco, pesquisadores passaram a perceber que muitos fenômenos sociais não podiam ser adequadamente explicados apenas pelas categorias tradicionais de classe, família ou instituição. As relações concretas estabelecidas entre indivíduos frequentemente ultrapassavam os limites formais dessas organizações, formando estruturas flexíveis de cooperação, influência e circulação de recursos. Era necessário, portanto, desenvolver instrumentos capazes de captar essas conexões informais que atravessavam diferentes grupos sociais (Mitchell, 1969).

Nesse contexto destaca-se o trabalho de John Arundel Barnes, responsável por introduzir, em 1954, uma das primeiras utilizações sistemáticas da expressão *social network*. Ao estudar uma comunidade pesqueira na Noruega, Barnes observou que os habitantes participavam simultaneamente de múltiplos sistemas de relacionamento — familiares, econômicos, religiosos, políticos e de amizade — formando uma trama complexa de vínculos que não podia ser reduzida às estruturas institucionais formais. Para descrever essa realidade, o autor empregou a metáfora da rede, entendendo-a como um conjunto de relações interpessoais que se sobrepunham continuamente e conectavam diferentes segmentos da comunidade. A importância desse estudo reside precisamente na transformação da rede

social em categoria analítica autônoma, capaz de explicar fenômenos que escapavam às classificações tradicionais da Sociologia (Barnes, 1954).

A metáfora da rede revelou-se particularmente fecunda porque permitia compreender a sociedade como uma estrutura dinâmica e aberta. Diferentemente das abordagens centradas em grupos rigidamente delimitados, a perspectiva das redes reconhecia que os indivíduos pertencem simultaneamente a diferentes circuitos de relacionamento. Uma pessoa pode integrar, ao mesmo tempo, uma família, uma equipe profissional, uma comunidade religiosa, um grupo esportivo, uma associação política e um conjunto de amizades construídas em ambientes digitais. Cada um desses vínculos influencia suas possibilidades de acesso à informação, de mobilização de recursos, de formação de opiniões e de participação social. A rede deixa de ser apenas uma imagem descritiva e passa a constituir uma forma específica de compreender a própria organização da vida coletiva (Barnes, 1954; Mitchell, 1969).

Outro avanço importante ocorreu com os estudos de Elizabeth Bott sobre relações familiares. Ao investigar casais residentes em Londres, Bott (1957) demonstrou que a intensidade e a densidade das redes externas influenciavam diretamente a divisão de papéis no interior da família. Casais inseridos em redes amplas e fortemente conectadas tendiam a apresentar formas distintas de organização doméstica em comparação com aqueles cujas relações externas eram mais restritas. Essa constatação ampliava significativamente o alcance da análise de redes, mostrando que características tradicionalmente atribuídas ao âmbito privado eram, na realidade, profundamente condicionadas pelas estruturas relacionais nas quais os indivíduos estavam inseridos.

Essas contribuições convergiram para uma mudança epistemológica de grande alcance. A sociedade passou gradativamente a ser concebida não apenas como um conjunto de instituições, normas ou categorias sociais, mas como uma configuração de relações. A posição ocupada pelos indivíduos dentro dessas configurações tornou-se elemento central para explicar oportunidades, constrangimentos, circulação de informações e distribuição de poder. Em vez de perguntar apenas “quem são os atores?”, os pesquisadores passaram a perguntar “como esses atores estão conectados?” e “quais consequências decorrem dessas conexões?”. Essa inversão metodológica constitui uma das bases da moderna Análise

de Redes Sociais e permanece fundamental para compreender, posteriormente, as plataformas digitais e seus efeitos sobre a circulação da informação (Wellman, 1988; Wasserman; Faust, 1994).

Nas décadas de 1960 e 1970, esse campo recebeu novo impulso com o desenvolvimento de métodos quantitativos, teoria dos grafos e recursos computacionais que permitiram representar matematicamente estruturas relacionais cada vez mais complexas. A Análise de Redes Sociais deixou de constituir apenas uma metáfora sociológica e passou a consolidar-se como um programa científico interdisciplinar, incorporando contribuições da Matemática, da Estatística, da Ciência da Computação, da Antropologia, da Sociologia e da Ciência Política. A noção de rede adquiria, assim, crescente precisão conceitual, permitindo investigar empiricamente fenômenos como difusão de informações, formação de comunidades, liderança, inovação, cooperação, influência e propagação de comportamentos coletivos (Wasserman; Faust, 1994).

Foi nesse contexto que Mark Granovetter apresentou uma das contribuições mais influentes da Sociologia contemporânea ao publicar, em 1973, o artigo *The Strength of Weak Ties*. Contrariando a expectativa intuitiva de que os vínculos mais importantes seriam aqueles estabelecidos com familiares e amigos íntimos, Granovetter (1973) demonstrou que os chamados “laços fracos” desempenham papel decisivo na circulação de novas informações entre diferentes grupos sociais. Relações menos intensas — como colegas de trabalho, conhecidos ou contatos ocasionais — frequentemente funcionam como pontes capazes de conectar comunidades distintas, ampliando o acesso a oportunidades, conhecimentos e recursos que dificilmente circulariam em grupos excessivamente fechados.

A teoria da força dos laços fracos tornou-se um dos pilares para compreender o funcionamento das redes contemporâneas, inclusive das plataformas digitais. Muito antes da existência do *Facebook* ou do *LinkedIn*, Granovetter (1973) já demonstrava que a arquitetura das conexões influencia profundamente os fluxos informacionais e a formação das relações sociais. A inovação das redes digitais não consistiria, portanto, na criação dessas dinâmicas, mas em sua amplificação exponencial por meio de tecnologias capazes de conectar milhões de indivíduos em tempo real. Essa constatação representa uma passagem fundamental para o argumento desenvolvido nesta tese: compreender as redes sociais contemporâ-

neas exige reconhecer que seus fundamentos sociológicos antecedem em muitas décadas a emergência da internet, ainda que as tecnologias digitais tenham transformado radicalmente sua escala, velocidade e capacidade de influência. (Wellman, 1988).

A consolidação da Análise de Redes Sociais (*Social Network Analysis* – SNA) ao longo das décadas de 1970 e 1980 representou um marco decisivo na compreensão científica das relações humanas. Até então, grande parte das investigações sociológicas concentrava-se nas características individuais dos sujeitos ou nas propriedades das instituições sociais. A perspectiva relacional introduziu uma mudança paradigmática ao demonstrar que muitos fenômenos coletivos somente podem ser explicados quando se analisa a estrutura das conexões entre os atores. Nessa abordagem, indivíduos, organizações e comunidades deixam de ser concebidos como unidades isoladas para serem compreendidos como nós inseridos em sistemas complexos de relações, cujas propriedades influenciam oportunidades, fluxos de informação, distribuição de recursos e exercício do poder. A rede passa, assim, a constituir simultaneamente objeto de investigação, método analítico e perspectiva epistemológica para compreender a organização social (Wellman, 1988; Wasserman; Faust, 1994).

Barry Wellman (1988) desempenhou papel central nesse processo ao propor que a unidade fundamental da análise sociológica não deveria ser o grupo social rigidamente delimitado, mas o conjunto de relações efetivamente estabelecidas entre indivíduos e instituições. Em diversos estudos, o autor demonstrou que as pessoas vivem inseridas em múltiplas redes parcialmente sobrepostas, compostas por vínculos familiares, profissionais, econômicos, religiosos, afetivos e comunitários. Essas redes apresentam diferentes níveis de densidade, reciprocidade, centralidade e intensidade, influenciando diretamente a circulação de informações, o apoio social, a construção de confiança e a formação de capital social. Ao deslocar o foco do pertencimento institucional para a conectividade relacional, Wellman contribuiu para consolidar uma visão segundo a qual a sociedade é continuamente produzida pelas conexões que seus integrantes estabelecem entre si.

Essa perspectiva permitiu superar uma dicotomia historicamente presente na Sociologia entre indivíduo e sociedade. Em vez de considerar que os sujeitos são determinados exclusivamente pelas estruturas sociais ou, em sentido oposto,

que a sociedade resulta apenas da soma das ações individuais, a Análise de Redes Sociais evidencia que ambos se constituem reciprocamente. Os indivíduos moldam suas redes ao estabelecer vínculos, mas também são moldados pelas oportunidades, restrições e fluxos informacionais produzidos por essas mesmas redes. A posição ocupada por cada ator influencia significativamente seu acesso ao conhecimento, sua capacidade de influência, seu potencial de inovação e suas possibilidades de mobilização coletiva. A estrutura relacional deixa de ser mero pano de fundo para tornar-se elemento constitutivo da própria ação social (Wellman, 1988; Wasserman; Faust, 1994).

A consolidação metodológica desse campo foi amplamente sistematizada por Stanley Wasserman e Katherine Faust na obra *Social Network Analysis: Methods and Applications*, publicada em 1994. Considerada uma das principais referências da área, a obra organizou conceitos, técnicas estatísticas e procedimentos analíticos que conferiram maior rigor científico aos estudos sobre redes. Nela, a rede social é compreendida como uma estrutura composta por atores — indivíduos, grupos, organizações ou instituições — conectados por diferentes tipos de relações, cujas propriedades podem ser representadas matematicamente por meio da teoria dos grafos. Essa formalização permitiu que conceitos como densidade, centralidade, proximidade, intermediação (*betweenness*), coesão e componentes estruturais passassem a ser mensurados empiricamente, ampliando consideravelmente o potencial explicativo da abordagem relacional.

A incorporação desses instrumentos quantitativos não reduziu a complexidade dos fenômenos sociais a modelos matemáticos simplificados. Ao contrário, permitiu compreender que determinadas propriedades emergem da própria configuração da rede e não podem ser explicadas exclusivamente pelas características individuais dos atores. Um indivíduo altamente qualificado pode exercer pouca influência caso ocupe uma posição periférica na estrutura relacional; por outro lado, sujeitos com poucos recursos individuais podem adquirir enorme capacidade de difusão de informações quando localizados em posições estratégicas de intermediação entre diferentes comunidades. Essa compreensão seria posteriormente decisiva para interpretar o funcionamento das plataformas digitais, nas quais algoritmos frequentemente reforçam ou reorganizam posições estruturais de visibilidade e influência (Wasserman; Faust, 1994; Recuero, 2018).

A partir desse momento, a noção de rede passou a transcender os limites da Sociologia. Economistas utilizaram a abordagem para compreender cadeias produtivas e mercados; cientistas políticos investigaram coalizões, circulação de poder e redes de governança; epidemiologistas examinaram padrões de disseminação de doenças; estudiosos da comunicação analisaram fluxos informacionais; pesquisadores da educação passaram a investigar comunidades de aprendizagem; e cientistas da computação incorporaram modelos relacionais para compreender arquiteturas digitais. A força explicativa da teoria das redes reside precisamente em sua capacidade de revelar padrões de organização comuns a fenômenos extremamente distintos, sem ignorar as especificidades históricas e culturais de cada contexto (Castells, 1999; Wasserman; Faust, 1994).

Entretanto, seria um equívoco interpretar essa expansão como mera universalização de uma metáfora. A literatura demonstra que diferentes tipos de redes apresentam propriedades distintas e produzem efeitos específicos. Redes fortemente densas tendem a favorecer confiança, cooperação e identidade coletiva, mas podem reduzir a circulação de informações novas devido à elevada redundância dos vínculos. Redes mais abertas e diversificadas, por sua vez, facilitam inovação, criatividade e acesso a recursos externos, embora frequentemente apresentem menor estabilidade relacional. Essa tensão entre coesão e abertura acompanha praticamente toda a literatura sobre redes sociais e permanece atual na análise das plataformas digitais contemporâneas, nas quais algoritmos podem simultaneamente ampliar conexões e reforçar processos de homofilia, segregação e formação de comunidades altamente fechadas (Granovetter, 1973; Wellman, 1988).

Outro aspecto fundamental desenvolvido pela literatura refere-se à relação entre redes e capital social. Embora o conceito possua diferentes formulações, autores como Pierre Bourdieu, James Coleman e Robert Putnam demonstraram que as relações sociais constituem recursos capazes de produzir benefícios individuais e coletivos. Informação, confiança, cooperação, apoio emocional, reconhecimento e oportunidades profissionais frequentemente circulam através das redes de relacionamento, configurando um patrimônio social cuja distribuição é profundamente desigual. O acesso às redes, portanto, não representa apenas uma dimensão da sociabilidade, mas um fator decisivo para compreender desigualda-

des econômicas, educacionais, culturais e políticas (Bourdieu, 1980; Coleman, 1988; Putnam, 2000).

Essa compreensão adquire especial relevância quando observada à luz da transformação tecnológica iniciada nas últimas décadas do século XX. Até esse momento, as redes sociais eram predominantemente analisadas como estruturas relacionais existentes entre sujeitos concretos, independentemente do suporte tecnológico utilizado para sua comunicação. A emergência da informática, das telecomunicações digitais e, posteriormente, da internet não altera a existência dessas redes, mas modifica profundamente as condições materiais pelas quais elas são estabelecidas, mantidas e ampliadas. O desenvolvimento tecnológico não substitui as relações sociais; ele reorganiza sua escala, velocidade, alcance geográfico, permanência e capacidade de armazenamento. A tecnologia passa a constituir parte integrante da infraestrutura sobre a qual as redes sociais se expandem (Castells, 1999).

É precisamente nesse ponto que emerge uma das contribuições mais influentes da Sociologia contemporânea: a teoria da sociedade em rede desenvolvida por Manuel Castells. Diferentemente de interpretações que atribuem à internet o surgimento das redes sociais, Castells (1999) argumenta que as redes sempre constituíram uma forma histórica de organização humana. O que distingue o final do século XX não é a invenção das redes, mas a convergência entre tecnologias digitais, globalização econômica, reorganização do capitalismo e revolução informacional, produzindo uma nova lógica estrutural para o funcionamento da sociedade. As redes deixam de ocupar posições periféricas para tornarem-se o princípio organizador dominante das relações econômicas, políticas, culturais e comunicacionais.

Na trilogia *A Era da Informação*, especialmente no primeiro volume, *A Sociedade em Rede*, publicado originalmente em 1996, Castells propõe que a principal transformação histórica daquele período consiste na emergência de uma nova morfologia social. Em vez de organizações rigidamente hierarquizadas, características da sociedade industrial, observa-se crescente predominância de estruturas flexíveis, descentralizadas e altamente conectadas. Empresas organizam cadeias produtivas globais; mercados financeiros operam em tempo real; universidades estabelecem cooperação internacional; movimentos sociais

articulam mobilizações transnacionais; governos compartilham informações em escala planetária. A lógica em rede torna-se elemento estruturante da organização contemporânea, ultrapassando amplamente o campo das tecnologias digitais.

Ao utilizar a expressão “sociedade em rede”, Castells (1999) não pretende afirmar que todas as formas anteriores de organização desapareceram. Instituições burocráticas, Estados nacionais, empresas hierarquizadas e organizações tradicionais continuam existindo. O argumento central consiste em demonstrar que essas instituições passam progressivamente a operar em interação permanente com redes informacionais distribuídas globalmente. A conectividade deixa de ser característica periférica para transformar-se em condição fundamental da produção econômica, da circulação cultural, da inovação científica e da comunicação social. As redes convertem-se na infraestrutura dominante da sociedade informacional.

Essa interpretação apresenta consequências importantes para a compreensão das redes sociais contemporâneas. Em primeiro lugar, evidencia que a internet não deve ser concebida como simples tecnologia de comunicação, mas como parte de uma transformação estrutural mais ampla que reorganiza processos produtivos, relações de poder, circulação de conhecimento e formas de interação social. Em segundo lugar, demonstra que o desenvolvimento das plataformas digitais não representa ruptura absoluta com o passado, mas aprofundamento de tendências relacionais já presentes nas formas históricas de organização humana. Finalmente, permite compreender que os fenômenos associados à circulação da informação — inclusive aqueles relacionados à desinformação e à pós-verdade — não decorrem exclusivamente das plataformas tecnológicas, mas da articulação entre estruturas sociais, interesses econômicos, arquiteturas técnicas e práticas culturais que caracterizam a sociedade em rede (Castells, 1999; Castells, 2003).

É justamente essa perspectiva que prepara a passagem para um novo estágio da evolução conceitual das redes sociais. Se a primeira etapa corresponde à compreensão das redes como estruturas relacionais investigadas pela Sociologia e a segunda refere-se à consolidação da sociedade em rede como paradigma organizador da contemporaneidade, a etapa seguinte será marcada pelo surgimento da internet comercial, da Web 2.0 e dos chamados *Social Ne-*

work Sites. Nesse momento, *as redes sociais deixam de ser apenas objeto de análise científica para tornar-se ambientes digitais de interação cotidiana*, inaugurando uma transformação profunda na produção, circulação e visibilidade das relações sociais, tema que será desenvolvido a seguir (Castells, 2003; Boyd; Elisson, 2007).

A passagem da sociedade em rede para o universo das plataformas digitais não ocorreu de maneira abrupta nem pode ser explicada exclusivamente pelo avanço tecnológico. Entre a consolidação da internet e o surgimento das atuais plataformas de interação social desenvolveu-se um conjunto de transformações técnicas, econômicas e culturais que alterou profundamente o modo como os indivíduos produzem, distribuem e consomem informações. A compreensão desse processo exige reconhecer que a internet, em sua configuração inicial, possuía características bastante distintas daquelas observadas nas plataformas contemporâneas. Durante as décadas de 1980 e 1990, predominavam páginas estáticas, ambientes institucionais, listas de discussão, fóruns temáticos, correio eletrônico e sistemas de navegação baseados em diretórios relativamente estáveis. Embora já existissem formas relevantes de interação entre usuários, a arquitetura da rede permanecia fortemente orientada para a disponibilização de conteúdos produzidos por organizações, universidades, empresas e instituições públicas. A participação dos usuários, embora presente, encontrava limitações técnicas que restringiam sua capacidade de produzir informação em larga escala (Castells, 2003; Lévy, 1999).

A expansão da internet comercial durante os anos 1990 produziu um crescimento exponencial do número de usuários conectados, mas foi apenas no início do século XXI que se consolidou uma mudança estrutural na lógica de funcionamento da própria web. Essa transformação ficou conhecida como Web 2.0, expressão popularizada por Tim O'Reilly em 2005 para descrever um conjunto de aplicações cujo principal diferencial consistia na valorização da participação ativa dos usuários. Mais do que uma atualização tecnológica, a Web 2.0 representou uma mudança de paradigma comunicacional. O usuário deixou de ocupar predominantemente a posição de receptor de conteúdos para assumir, simultaneamente, os papéis de produtor, distribuidor, comentarista e curador de informações. Essa alteração modificou profundamente as formas de circulação do co-

nhecimento, ampliando a capacidade de produção colaborativa e descentralizada de conteúdos (O'Reilly, 2005; Lévy, 1999).

Pierre Lévy (1999) antecipava parte dessa transformação ao desenvolver o conceito de inteligência coletiva. Segundo o autor, o ambiente digital favoreceria novas formas de produção compartilhada do conhecimento, nas quais diferentes indivíduos poderiam contribuir simultaneamente para a construção de soluções, informações e aprendizagens. A internet seria capaz de potencializar competências distribuídas socialmente, permitindo que o conhecimento deixasse de depender exclusivamente de instituições centralizadas e passasse a emergir de processos colaborativos entre sujeitos conectados. A inteligência coletiva, nessa perspectiva, não elimina a importância do conhecimento especializado, mas amplia as possibilidades de cooperação entre diferentes produtores de informação, inaugurando formas inéditas de criação e circulação do saber.

A emergência da Web 2.0 intensificou significativamente essas possibilidades colaborativas. Blogs, wikis, sistemas de compartilhamento de fotografias, plataformas de vídeo e ambientes de interação passaram a oferecer interfaces intuitivas, permitindo que pessoas sem conhecimentos avançados de programação produzissem conteúdos acessíveis globalmente. Essa democratização técnica reduziu custos de publicação, ampliou a diversidade de vozes presentes no espaço público digital e modificou profundamente a ecologia da comunicação. Pela primeira vez, milhões de indivíduos puderam tornar públicas suas opiniões, experiências e produções culturais sem depender da mediação editorial tradicional exercida por jornais, emissoras de televisão, editoras ou instituições acadêmicas (Castells, 2003; Lévy, 1999).

Entretanto, a ampliação da participação dos usuários não significou a eliminação dos processos de mediação. Ao contrário, novas formas de controle passaram gradativamente a substituir aquelas exercidas pelos meios de comunicação tradicionais. Se anteriormente a seleção dos conteúdos dependia principalmente de decisões editoriais humanas, no ambiente digital essa função começou a ser compartilhada por mecanismos automáticos de classificação, indexação e recomendação. A arquitetura técnica das plataformas passou a desempenhar papel crescente na definição do que seria visível, compartilhável e potencialmente relevante para diferentes públicos. Assim, a descentralização da produção informa-

cional coexistiu com novas formas de centralização da circulação dos conteúdos (Gillespie, 2018; Van Dijck, 2013).

Foi nesse contexto que surgiram os chamados *Social Network Sites* (SNS), expressão cuja formulação conceitual foi sistematizada por Danah Boyd e Nicole Ellison em um artigo publicado, em 2007, no *Journal of Computer-Mediated Communication*. As autoras observam que, embora diversas plataformas digitais fossem frequentemente denominadas “redes sociais”, havia pouca precisão terminológica sobre aquilo que efetivamente caracterizava esse tipo específico de ambiente digital. Seu objetivo consistiu justamente em estabelecer uma definição conceitualmente rigorosa que permitisse distinguir o fenômeno tecnológico das estruturas sociais estudadas pela Sociologia desde o início do século XX.

Segundo Boyd e Ellison (2007), um *Social Network Site* pode ser definido como um serviço baseado na internet que permite aos usuários construir um perfil público ou semipúblico dentro de um sistema delimitado, articular uma lista de outros usuários com os quais estabelecem conexões e visualizar ou percorrer tanto sua própria lista de relações quanto aquelas estabelecidas por outros participantes da plataforma. Essa definição tornou-se referência internacional por deslocar o foco da tecnologia para as funcionalidades específicas desses ambientes digitais. O elemento distintivo dos sites de redes sociais não reside apenas na possibilidade de comunicação, mas na visibilidade das conexões sociais e na capacidade de tornar públicas, navegáveis e persistentes as relações estabelecidas entre seus usuários.

Essa formulação introduz uma distinção conceitual extremamente relevante para a pesquisa científica. A expressão “rede social” refere-se, propriamente, ao conjunto de relações estabelecidas entre atores sociais. Já os *Social Network Sites* constituem plataformas tecnológicas projetadas para dar suporte, registrar, ampliar e tornar visíveis essas relações. Em outras palavras, *Facebook*, *Instagram*, *LinkedIn* ou *TikTok* não são, em sentido estrito, as redes sociais propriamente ditas; são ambientes digitais que permitem a manifestação, reorganização e expansão de redes sociais previamente existentes ou construídas durante a interação online. Confundir ambos os conceitos significa reduzir um fenômeno sociológico complexo a uma tecnologia específica, obscurecendo a historicidade das relações sociais que antecederam amplamente o surgimento da internet (Boyd; Ellison, 2007; Recuero, 2018).

Essa diferenciação foi amplamente desenvolvida, no contexto brasileiro, por Raquel Recuero (2018), cuja produção constitui uma das principais referências nacionais sobre redes sociais na internet. Para a autora, as redes sociais devem ser compreendidas como estruturas compostas por atores e conexões, sendo os atores representados por indivíduos, grupos ou organizações e as conexões constituídas pelas interações e pelos laços sociais estabelecidos entre eles. As plataformas digitais funcionam como espaços nos quais essas conexões tornam-se observáveis, registráveis e analisáveis, permitindo que fenômenos tradicionalmente invisíveis sejam identificados por meio da circulação de informações, da formação de comunidades e da dinâmica das interações.

Recuero (2018) enfatiza que a internet não cria automaticamente relações sociais significativas. A existência de uma conexão técnica entre perfis não implica, necessariamente, a formação de um vínculo social consistente. Curtidas, compartilhamentos, comentários, mensagens privadas e interações frequentes podem assumir significados bastante distintos conforme o contexto relacional em que ocorrem. Algumas conexões permanecem superficiais e episódicas; outras desenvolvem vínculos afetivos duradouros capazes de produzir confiança, cooperação e capital social. Essa distinção é particularmente importante porque impede interpretações quantitativas simplificadas que associam número de seguidores ou de contatos ao efetivo fortalecimento das relações sociais.

Outro aspecto desenvolvido pela autora refere-se à permanência dos rastros digitais. Diferentemente das interações presenciais, muitas relações estabelecidas nas plataformas digitais deixam registros relativamente permanentes, permitindo que comentários, compartilhamentos, fotografias e conexões permaneçam disponíveis durante longos períodos. Essa característica amplia significativamente as possibilidades analíticas da pesquisa científica, pois torna observáveis processos relacionais que anteriormente permaneciam invisíveis à investigação empírica. Ao mesmo tempo, introduz novas formas de vigilância, monitoramento e produção de dados que ultrapassam amplamente os objetivos originais da comunicação interpessoal (Recuero, 2018).

A consolidação dos *Social Network Sites* também provocou uma crescente confusão terminológica no debate público. Expressões como “rede social”, “mídia social”, “plataforma digital” e “rede social digital” passaram a ser fre-

quentemente utilizadas como sinônimos, embora designem fenômenos distintos. Tal imprecisão não constitui mero problema semântico, mas compromete a compreensão científica das transformações contemporâneas. Diferenciar esses conceitos torna-se condição indispensável para analisar adequadamente os efeitos produzidos pela circulação da informação, pela arquitetura algorítmica e pela emergência da pós-verdade.

Em sentido sociológico, conforme discutido anteriormente, a rede social corresponde à estrutura de relações estabelecidas entre atores sociais. Ela pode existir independentemente de qualquer tecnologia específica, manifestando-se em grupos familiares, comunidades religiosas, organizações profissionais, movimentos sociais ou relações de amizade. A mídia social, por sua vez, refere-se ao conjunto de tecnologias digitais destinadas à produção, publicação, compartilhamento e circulação de conteúdos gerados pelos próprios usuários. Enquanto a rede social diz respeito às relações, a mídia social refere-se aos meios técnicos utilizados para comunicar, distribuir e tornar públicos esses conteúdos (Kaplan; Haenlein, 2010).

Os *Social Network Sites* constituem uma categoria específica dentro do universo mais amplo das mídias sociais. Toda plataforma classificada como site de rede social é também uma mídia social, mas nem toda mídia social corresponde a um site de rede social. Plataformas como *Wikipédia*, *YouTube* ou *blogs* podem favorecer intensa produção colaborativa de conteúdos sem necessariamente organizar suas funcionalidades prioritariamente em torno da visualização das conexões sociais entre usuários. Os sites de redes sociais distinguem-se justamente por estruturar sua arquitetura em torno dos perfis individuais, das listas de conexões e da navegabilidade das relações sociais estabelecidas entre seus participantes (Boyd; Ellison, 2007).

Essa diferenciação prepara uma etapa ainda mais complexa da evolução conceitual. À medida que empresas como *Meta*, *Google*, *ByteDance*, *Microsoft* e outras passaram a integrar diferentes serviços digitais em ecossistemas tecnológicos altamente conectados, os antigos sites de redes sociais transformaram-se progressivamente em plataformas digitais multifuncionais. Já não se limitam a mediar relações interpessoais ou facilitar o compartilhamento de conteúdos. Tornaram-se infraestruturas responsáveis por organizar fluxos globais de informa-

ção, coletar dados comportamentais em larga escala, modular padrões de atenção, selecionar conteúdos por meio de algoritmos e produzir novas formas de governança informacional. Compreender essa transformação exige ultrapassar definitivamente a ideia de redes sociais como simples espaços de interação e analisá-las como infraestruturas sociotécnicas complexas, discussão que será desenvolvida a seguir (Van Dijck, 2013; Gillespie, 2018).

A transformação dos sites de redes sociais em plataformas digitais constitui uma das mudanças mais profundas da ecologia informacional contemporânea. Embora os primeiros ambientes de interação online tenham sido concebidos principalmente como espaços destinados ao encontro entre pessoas e ao compartilhamento de conteúdos, o desenvolvimento econômico e tecnológico das grandes empresas do setor modificou substancialmente essa finalidade. Progressivamente, essas organizações deixaram de oferecer apenas serviços de comunicação para operar como infraestruturas digitais capazes de articular múltiplos sistemas de informação, publicidade, comércio eletrônico, entretenimento, inteligência artificial, geolocalização e processamento massivo de dados. Essa mudança exige uma revisão conceitual importante: compreender *Facebook*, *Instagram*, *TikTok*, *X* ou *LinkedIn* apenas como “redes sociais” revela-se insuficiente para explicar sua complexidade funcional e seu impacto sobre a organização da vida social contemporânea (Van Dijck, 2013; Srnicek, 2017).

José Van Dijck (2013) foi uma das principais autoras a sistematizar essa mudança de perspectiva ao analisar o processo de plataformização da sociedade. Em *The Culture of Connectivity*, publicado em 2013, a autora argumenta que plataformas digitais não podem ser compreendidas apenas como instrumentos tecnológicos neutros colocados à disposição dos usuários. Elas constituem sistemas sociotécnicos complexos nos quais arquitetura computacional, modelos econômicos, interesses corporativos, formas de governança e práticas culturais encontram-se profundamente articulados. A conectividade, nesse contexto, deixa de representar apenas a possibilidade técnica de estabelecer relações entre indivíduos e transforma-se em um modelo econômico baseado na coleta, organização, análise e monetização permanente das interações produzidas pelos usuários.

Essa interpretação rompe com uma narrativa amplamente difundida durante os primeiros anos da internet, segundo a qual as plataformas digitais seriam

ambientes essencialmente democráticos e horizontalizados, capazes de eliminar antigos mecanismos de controle da comunicação. Van Dijck (2013) demonstra que, embora essas tecnologias tenham ampliado significativamente as possibilidades de participação pública, elas também instituíram novas formas de mediação extremamente sofisticadas. A interação entre usuários ocorre sempre dentro de arquiteturas computacionais cuidadosamente projetadas, cujas regras de funcionamento permanecem, em grande medida, invisíveis para aqueles que delas participam. As escolhas realizadas pelos usuários são constantemente orientadas por interfaces, mecanismos de recomendação, sistemas de classificação, critérios de visibilidade e formas específicas de organização da informação definidas pelas próprias plataformas.

Essa constatação conduz a uma alteração significativa no próprio conceito de mediação. Nos meios de comunicação tradicionais, a mediação era exercida predominantemente por jornalistas, editores, produtores culturais e instituições responsáveis por selecionar conteúdos considerados relevantes para publicação. Nas plataformas digitais, parte importante dessa função passa a ser desempenhada por sistemas algorítmicos capazes de processar enormes volumes de dados em tempo real. Embora decisões humanas continuem presentes na programação, no treinamento e na supervisão desses sistemas, a seleção cotidiana daquilo que será exibido para cada usuário depende crescentemente de processos computacionais automatizados. A mediação editorial transforma-se, assim, em mediação algorítmica (Gillespie, 2018; Van Dijck, 2013).

Tarleton Gillespie (2018) aprofunda essa discussão ao demonstrar que os algoritmos não podem ser compreendidos como simples procedimentos matemáticos neutros. Segundo o autor, eles exercem papel decisivo na organização da visibilidade pública contemporânea. Ao selecionar quais conteúdos aparecerão primeiro, quais permanecerão ocultos, quais serão recomendados e quais serão considerados relevantes, os algoritmos influenciam diretamente aquilo que os indivíduos percebem como realidade social. Ainda que não determinem completamente o comportamento humano, participam ativamente da construção das agendas informacionais, das prioridades discursivas e das oportunidades de circulação de diferentes narrativas. Desse modo, tornam-se atores centrais na organização do espaço público digital.

É importante destacar que os algoritmos não operam em um vazio social. Eles são desenvolvidos por equipes técnicas vinculadas a empresas privadas, treinados sobre conjuntos específicos de dados, submetidos a objetivos econômicos e continuamente ajustados conforme métricas de desempenho definidas institucionalmente. Sua aparente objetividade decorre, muitas vezes, da invisibilidade dos critérios utilizados para organizar os fluxos informacionais. A classificação de conteúdos como relevantes, interessantes ou prioritários não resulta apenas da aplicação automática de fórmulas matemáticas, mas incorpora escolhas políticas, comerciais e culturais inscritas no próprio desenho das plataformas. A opacidade desses processos dificulta o escrutínio público e amplia o poder das empresas responsáveis por sua administração (Gillespie, 2018).

Essa reorganização tecnológica coincide com profundas transformações econômicas analisadas por Nick Srnicek (2017) ao desenvolver o conceito de capitalismo de plataforma. Segundo o autor, as plataformas digitais representam um novo modelo de organização econômica caracterizado pela centralidade dos dados como principal recurso estratégico. Diferentemente das empresas industriais clássicas, cuja riqueza dependia predominantemente da produção material de bens, as plataformas baseiam seu funcionamento na captura, processamento e comercialização de informações produzidas continuamente pelas atividades de seus usuários. Cada curtida, compartilhamento, pesquisa, fotografia, comentário, localização geográfica ou tempo de permanência transforma-se em dado potencialmente explorável para fins econômicos.

A importância dos dados nesse novo modelo ultrapassa sua utilização para fins publicitários. Eles constituem matéria-prima indispensável para o desenvolvimento de sistemas de recomendação, personalização de conteúdos, reconhecimento de padrões comportamentais e treinamento de modelos computacionais cada vez mais sofisticados. Quanto maior o volume de informações coletadas, maior tende a ser a capacidade das plataformas de prever preferências, antecipar comportamentos e oferecer conteúdos personalizados aos seus usuários. Estabelece-se, assim, um ciclo cumulativo no qual dados produzem melhores serviços, que atraem novos usuários, que geram mais dados, fortalecendo progressivamente a posição competitiva das grandes plataformas digitais (Srnicek, 2017).

Essa dinâmica econômica altera profundamente o funcionamento das próprias redes sociais digitais. Nos primeiros anos da internet, as conexões

estabelecidas entre usuários constituíam o principal elemento organizador das plataformas. Progressivamente, entretanto, as relações interpessoais passaram a coexistir com complexos sistemas de processamento algorítmico capazes de selecionar conteúdos independentemente da proximidade relacional entre seus produtores e receptores. O que aparece na tela do usuário já não depende exclusivamente das pessoas que ele decidiu seguir ou adicionar como contato, mas de cálculos probabilísticos realizados continuamente pelos sistemas computacionais. As plataformas deixam de refletir apenas as redes sociais existentes e passam a participar ativamente da produção das próprias experiências informacionais vivenciadas pelos indivíduos (Van Dijck, 2013; Gillespie, 2018).

Uma consequência direta desse processo é a crescente personalização da experiência digital. Dois usuários conectados à mesma plataforma, realizando pesquisas semelhantes ou pertencendo ao mesmo grupo social, podem receber conjuntos bastante diferentes de informações, recomendações e conteúdos. Essa diferenciação decorre da combinação entre histórico de navegação, preferências inferidas, interações anteriores, localização geográfica, características demográficas e inúmeras outras variáveis processadas automaticamente pelos algoritmos. A experiência informacional deixa de ser predominantemente coletiva para assumir caráter progressivamente individualizado, produzindo ambientes comunicacionais altamente personalizados (Pariser, 2012; Van Dijck, 2013).

Eli Pariser (2012) denominou esse fenômeno de “bolha dos filtros” (*filter bubble*), argumentando que os sistemas de personalização tendem a apresentar aos usuários conteúdos compatíveis com seus interesses e comportamentos anteriores, reduzindo a exposição a perspectivas divergentes. Embora estudos posteriores tenham demonstrado que o funcionamento das plataformas é mais complexo do que a formulação inicial sugeria, a hipótese das bolhas informacionais permanece relevante ao evidenciar que os processos algorítmicos não distribuem informações de maneira homogênea entre todos os participantes da rede. Ao contrário, produzem ambientes comunicacionais diferenciados, nos quais determinados conteúdos tornam-se altamente visíveis para alguns grupos e praticamente inexistentes para outros.

Esse mecanismo articula-se estreitamente ao conceito de homofilia, amplamente estudado pela Sociologia das redes. A homofilia corresponde à ten-

dência de indivíduos estabelecerem relações preferencialmente com pessoas que compartilham características semelhantes, tais como valores, crenças, posição política, nível educacional, profissão ou interesses culturais. Nas plataformas digitais, essa disposição relacional pode ser reforçada pelos sistemas algorítmicos de recomendação, que identificam padrões de comportamento e passam a sugerir novos conteúdos, grupos e perfis compatíveis com preferências anteriormente manifestadas. O resultado não é apenas a ampliação das conexões existentes, mas a intensificação de determinados padrões de associação social (Mcpherson; Smith-lovin; Cook, 2001; Recuero, 2018).

Essa reorganização da conectividade produz efeitos que ultrapassam significativamente o âmbito da comunicação interpessoal. As plataformas tornam-se espaços de formação de opinião pública, mobilização política, construção de identidades coletivas, circulação de conhecimento científico, organização de movimentos sociais, produção cultural e disputas econômicas. Ao mesmo tempo em que ampliam possibilidades inéditas de participação democrática e colaboração em escala global, também concentram enorme capacidade de influenciar os fluxos informacionais que estruturam a vida social contemporânea. A arquitetura técnica das plataformas passa a desempenhar papel semelhante ao exercido historicamente pelas grandes instituições mediadoras da informação, embora submetida a lógicas econômicas e mecanismos de governança substancialmente distintos (Castells, 2009; Van Dijck, 2013).

É precisamente nesse cenário que a coleta de dados adquire importância estratégica não apenas para fins comerciais, mas também para compreender a reorganização das relações de poder na sociedade digital. As plataformas deixam de observar passivamente os comportamentos de seus usuários e passam a estruturar ambientes capazes de influenciar padrões de atenção, permanência, interação e engajamento. A produção de dados não constitui um efeito secundário da utilização desses serviços; ela representa uma das principais finalidades econômicas de sua arquitetura. Cada interação realizada pelos usuários alimenta continuamente sistemas destinados a aperfeiçoar modelos de previsão comportamental, personalização e recomendação de conteúdos.

Essa transformação prepara o terreno para uma das interpretações mais influentes da contemporaneidade: a teoria do capitalismo de vigilância desen-

volvida por Shoshana Zuboff. Ao deslocar a análise da simples coleta de dados para os mecanismos de previsão e modulação do comportamento humano, Zuboff (2019) demonstra que as plataformas digitais inauguram uma nova racionalidade econômica, na qual a experiência cotidiana dos indivíduos converte-se em matéria-prima para processos contínuos de monitoramento, análise e intervenção comportamental. É essa discussão que permitirá compreender, na sequência destas considerações, por que as plataformas algorítmicas se tornaram um dos principais ambientes de produção, amplificação e circulação da pós-verdade na sociedade contemporânea.

A compreensão das plataformas digitais como infraestruturas sociotécnicas alcança novo nível de profundidade com a teoria do capitalismo de vigilância desenvolvida por Shoshana Zuboff. Em *The Age of Surveillance Capitalism*, publicado em 2019, a autora sustenta que as grandes empresas de tecnologia inauguraram uma nova lógica de acumulação econômica baseada na apropriação sistemática da experiência humana como matéria-prima para processos de extração, processamento e comercialização de dados. Diferentemente dos modelos tradicionais de produção industrial, nos quais o valor econômico era gerado principalmente pela transformação de recursos materiais em mercadorias, o capitalismo de vigilância fundamenta-se na conversão da própria vida cotidiana em informação passível de tratamento computacional. Navegação na internet, deslocamentos geográficos, padrões de consumo, interações sociais, expressões emocionais, preferências culturais e comportamentos comunicacionais transformam-se continuamente em dados destinados à produção de previsões comportamentais economicamente exploráveis.

A originalidade da interpretação de Zuboff (2019) reside em demonstrar que o objetivo das plataformas digitais não consiste apenas em registrar comportamentos passados, mas sobretudo em antecipar comportamentos futuros. A coleta massiva de dados possibilita identificar padrões, probabilidades e regularidades estatísticas capazes de prever, com precisão crescente, decisões de consumo, interesses informacionais, deslocamentos, vínculos sociais e respostas emocionais. Esses modelos preditivos tornam-se produtos altamente valorizados por anunciantes, organizações políticas, empresas e diferentes agentes interessados em direcionar mensagens específicas para públicos cuidadosamente segmentados. Os

dados deixam de representar simples registros documentais da atividade humana para converterem-se em ativos estratégicos cuja principal finalidade consiste em produzir conhecimento probabilístico sobre comportamentos futuros.

Entretanto, a autora argumenta que a lógica do capitalismo de vigilância ultrapassa a mera previsão estatística. Progressivamente, as plataformas desenvolvem mecanismos destinados não apenas a antecipar ações humanas, mas também a influenciá-las. Sistemas de recomendação, notificações automáticas, organização visual das interfaces, sugestões de amizade, priorização de determinados conteúdos, reprodução automática de vídeos e inúmeras outras funcionalidades passam a integrar arquiteturas cuidadosamente planejadas para aumentar o tempo de permanência dos usuários, estimular determinadas formas de interação e orientar padrões específicos de comportamento. A fronteira entre observar e intervir torna-se progressivamente mais tênue. As plataformas deixam de atuar apenas como instrumentos de comunicação para participar ativamente da organização das experiências cotidianas de seus usuários (Zuboff, 2019).

Essa perspectiva aproxima-se das discussões contemporâneas sobre dataficação (*datafication*), conceito amplamente utilizado para descrever o processo pelo qual aspectos diversos da vida social passam a ser convertidos em dados digitais passíveis de armazenamento, processamento e análise. A dataficação não corresponde apenas ao aumento quantitativo da produção de informações digitais. Trata-se de uma transformação epistemológica por meio da qual fenômenos anteriormente compreendidos em termos qualitativos passam a ser representados como conjuntos de dados mensuráveis. Emoções, preferências, deslocamentos, relações sociais, hábitos de leitura, padrões de consumo e interações comunicacionais convertem-se em registros computacionais continuamente produzidos pelas atividades cotidianas dos indivíduos. A realidade social passa, assim, a ser progressivamente interpretada através da lógica dos dados (Van Dijck, 2014).

A produção contínua desses registros modifica profundamente as formas pelas quais o conhecimento sobre a sociedade é construído. Tradicionalmente, pesquisas sociológicas dependiam de entrevistas, observações, questionários ou levantamentos estatísticos realizados em períodos específicos. Nas plataformas digitais, ao contrário, milhões de dados são produzidos automaticamente a cada segundo, permitindo acompanhar em tempo real transformações comportamen-

tais de enorme complexidade. Essa capacidade analítica oferece oportunidades inéditas para investigação científica, planejamento urbano, monitoramento epidemiológico e desenvolvimento tecnológico. Simultaneamente, amplia significativamente o poder das organizações que controlam tais infraestruturas informacionais, produzindo novas assimetrias entre aqueles que geram os dados e aqueles que possuem condições de interpretá-los e explorá-los economicamente (Van Dijck, 2014; Zuboff, 2019).

Nesse cenário, a economia da atenção emerge como um dos principais fundamentos do funcionamento das plataformas digitais. Ainda no final do século XX, Herbert Simon (1971) observava que a abundância de informações produz inevitavelmente escassez de atenção. Quando o volume de conteúdos disponíveis supera a capacidade humana de processá-los, a atenção torna-se recurso limitado e altamente valioso. As plataformas digitais estruturam grande parte de sua arquitetura precisamente em torno da disputa por esse recurso escasso. Quanto maior o tempo de permanência do usuário, maior tende a ser a quantidade de dados produzidos, de anúncios visualizados e de oportunidades comerciais geradas. A atenção converte-se, assim, em importante ativo econômico da sociedade digital (Davenport; Beck, 2001).

A disputa pela atenção influencia diretamente o desenho das interfaces digitais. Elementos visuais, sistemas de notificações, mecanismos de atualização contínua, reprodução automática de conteúdos, recompensas variáveis e estratégias de personalização são organizados de modo a favorecer níveis elevados de engajamento. Embora tais recursos frequentemente ampliem a usabilidade e facilitem a navegação, também participam da construção de ambientes comunicacionais caracterizados pela permanente competição entre conteúdos. Nesse contexto, informações capazes de despertar surpresa, indignação, medo, humor ou intensa mobilização emocional tendem a apresentar maior potencial de circulação, uma vez que capturam mais rapidamente a atenção dos usuários e estimulam maiores índices de interação (Gillespie, 2018; Zuboff, 2019).

Essa característica possui implicações particularmente relevantes para compreender a circulação contemporânea da informação. As plataformas não selecionam necessariamente conteúdos verdadeiros ou falsos, científicos ou pseudocientíficos, responsáveis ou sensacionalistas. Seu funcionamento orienta-se

prioritariamente por indicadores relacionados à probabilidade de engajamento. Embora diferentes empresas tenham implementado mecanismos de moderação, verificação de fatos e redução da circulação de determinados conteúdos nocivos, a lógica econômica baseada na maximização da atenção continua exercendo influência significativa sobre os critérios de visibilidade informacional. A arquitetura técnica não determina automaticamente o conteúdo das mensagens, mas estabelece condições específicas para sua circulação e amplificação (Van Dijck, 2013; Gillespie, 2018).

É precisamente nesse ponto que se evidencia a importância da distinção entre neutralidade tecnológica e mediação sociotécnica. Durante muitos anos, difundiu-se a ideia de que algoritmos seriam instrumentos objetivos capazes de organizar informações segundo critérios puramente matemáticos. A literatura contemporânea contesta essa interpretação ao demonstrar que algoritmos incorporam decisões humanas em todas as etapas de seu desenvolvimento. A seleção dos dados utilizados para treinamento, a definição dos objetivos de otimização, os critérios de relevância, os parâmetros de classificação e os mecanismos de avaliação de desempenho resultam de escolhas realizadas por equipes técnicas e instituições responsáveis pela administração das plataformas. A automatização dos processos não elimina a dimensão social das decisões; apenas desloca parte delas para níveis menos visíveis da infraestrutura tecnológica (Gillespie, 2018; Kitchin, 2017).

Rob Kitchin (2017) observa que algoritmos devem ser compreendidos como sistemas sociotécnicos cuja atuação depende simultaneamente de modelos matemáticos, bases de dados, capacidades computacionais, contextos institucionais e interesses organizacionais. Sua eficácia não decorre exclusivamente da sofisticação técnica, mas da interação permanente entre elementos tecnológicos e sociais. Essa perspectiva impede interpretações deterministas segundo as quais os algoritmos exerceriam controle absoluto sobre o comportamento humano. Os usuários continuam produzindo interpretações, resistências, apropriações criativas e formas diversas de utilização das plataformas. Entretanto, tais práticas ocorrem dentro de ambientes cuidadosamente estruturados para orientar determinadas possibilidades de ação em detrimento de outras.

Essa articulação entre arquitetura tecnológica e ação humana aproxima-se da noção de *affordances*, frequentemente utilizada para compreender as possi-

bilidades de interação oferecidas pelos ambientes digitais. As plataformas não determinam comportamentos específicos, mas criam condições que tornam determinadas ações mais prováveis, mais fáceis ou mais visíveis do que outras. A existência de botões de compartilhamento instantâneo, sistemas de reação emocional, métricas públicas de popularidade, recomendações automáticas e notificações constantes constitui exemplos de *affordances* que influenciam significativamente a dinâmica comunicacional sem eliminar completamente a agência dos usuários. As tecnologias participam da construção das práticas sociais ao mesmo tempo em que são continuamente reinterpretadas e apropriadas pelos sujeitos que delas fazem uso (Boyd, 2011).

Outro aspecto decisivo refere-se à crescente opacidade dos sistemas algorítmicos. À medida que modelos computacionais se tornam mais complexos, reduz-se a capacidade dos próprios usuários de compreender os critérios pelos quais conteúdos lhes são apresentados. A experiência cotidiana nas plataformas caracteriza-se por elevada assimetria informacional: milhões de pessoas interagem continuamente com sistemas cujo funcionamento permanece amplamente desconhecido. Essa opacidade dificulta o exercício do controle democrático sobre infraestruturas que desempenham papel cada vez mais relevante na circulação de informações, na formação da opinião pública e na organização do debate político contemporâneo (Pasquale, 2015).

A centralização da capacidade de processamento de dados nas mãos de um número relativamente reduzido de empresas privadas constitui outra característica marcante desse processo. Embora a internet tenha sido frequentemente celebrada como ambiente descentralizado, observa-se progressiva concentração de funções essenciais em grandes corporações responsáveis por sistemas de busca, redes sociais, hospedagem de conteúdos, computação em nuvem, publicidade digital e inteligência artificial. Essa concentração amplia significativamente o poder dessas organizações para definir padrões tecnológicos, estabelecer regras de moderação, influenciar mercados informacionais e determinar as condições pelas quais bilhões de indivíduos acessam conhecimentos, estabelecem relações e participam da vida pública (Srnicek, 2017; Van Dijck, 2013).

Essas transformações não conduzem necessariamente a uma interpretação pessimista ou tecnofóbica das plataformas digitais. Os mesmos ambientes que

possibilitam intensa circulação de desinformação também ampliam oportunidades de divulgação científica, educação aberta, mobilização cidadã, organização comunitária, participação política e democratização do acesso à informação. Movimentos sociais historicamente marginalizados encontraram nas plataformas importantes instrumentos de visibilidade pública, denúncia de violações de direitos e articulação internacional. Instituições científicas passaram a comunicar diretamente resultados de pesquisas para públicos amplos. Processos educativos incorporaram novas possibilidades de aprendizagem colaborativa. O potencial emancipatório dessas tecnologias permanece significativo e não pode ser obscurecido por análises exclusivamente centradas em seus riscos (Castells, 2009; Lévy, 1999).

Entretanto, exatamente por exercerem papel central na organização da comunicação contemporânea, essas plataformas tornaram-se também espaços privilegiados para disputas em torno da produção da verdade pública. A combinação entre personalização algorítmica, economia da atenção, dataficação, modulação comportamental e concentração informacional cria condições particularmente favoráveis para a circulação acelerada de conteúdos capazes de mobilizar emoções, reforçar identidades coletivas e ampliar processos de polarização social. As plataformas não produzem automaticamente a pós-verdade, mas oferecem uma infraestrutura técnica e econômica extremamente eficiente para sua difusão, aspecto que será aprofundado na sequência deste assunto, quando se discutirá a relação entre arquitetura algorítmica, circulação informacional e construção contemporânea da realidade social.

As transformações analisadas até este ponto permitem compreender que as plataformas digitais não constituem apenas ambientes destinados ao compartilhamento de informações, mas ecossistemas comunicacionais cuja arquitetura influencia diretamente a forma pela qual os sujeitos percebem, interpretam e organizam a realidade social. Essa constatação conduz à questão central da presente investigação: *de que maneira a evolução das redes sociais para plataformas algorítmicas cria condições favoráveis à intensificação dos fenômenos associados à pós-verdade?* A resposta exige abandonar interpretações simplificadoras que atribuem às tecnologias digitais responsabilidade exclusiva pela circulação da desinformação. A literatura contemporânea demonstra que a relação entre plataformas e pós-verdade resulta da interação entre fatores tecnológicos, econômicos,

psicológicos, culturais e políticos, cuja convergência reorganiza profundamente as formas de produção e validação do conhecimento na sociedade contemporânea (Mcintyre, 2018; Cesarino, 2021).

A primeira condição diz respeito à reorganização da visibilidade informacional. Durante grande parte do século XX, a circulação pública de notícias dependia da atuação de instituições mediadoras relativamente bem definidas, como jornais, emissoras de rádio, televisão e editoras. Embora essas organizações jamais tenham sido neutras, seus conteúdos eram submetidos a procedimentos editoriais relativamente estáveis, envolvendo seleção de pautas, verificação de informações, revisão técnica e responsabilidade jurídica sobre aquilo que era publicado. A emergência das plataformas digitais descentralizou significativamente a produção da informação, permitindo que qualquer indivíduo passasse a ocupar simultaneamente as posições de produtor, distribuidor e consumidor de conteúdos. Essa ampliação da participação pública representa importante conquista democrática, mas também reduz a capacidade dos antigos mecanismos institucionais de filtragem e verificação prévia das informações (Castells, 2009; Van Dijck, 2013).

Entretanto, a descentralização da produção não implicou ausência de mediação. Conforme discutido anteriormente, a mediação editorial foi progressivamente substituída — ou complementada — pela mediação algorítmica. Os conteúdos que alcançam maior visibilidade não são necessariamente aqueles considerados mais relevantes sob critérios científicos, jornalísticos ou educacionais, mas aqueles que apresentam maior probabilidade de gerar interação, compartilhamento e permanência dos usuários na plataforma. Dessa forma, critérios tradicionalmente associados à qualidade informacional coexistem com métricas relacionadas ao engajamento, produzindo uma reorganização significativa da esfera pública digital. A arquitetura das plataformas passa a privilegiar a circulação de conteúdos potencialmente mais atrativos para os mecanismos algorítmicos, ainda que tais conteúdos não correspondam às informações mais rigorosas ou socialmente relevantes (Gillespie, 2018; Srnicek, 2017).

Essa reorganização torna-se particularmente significativa quando articulada às disposições cognitivas e emocionais próprias da experiência humana. Pesquisas nas áreas da Psicologia Cognitiva e das Ciências Sociais demons-

tram que os indivíduos não processam informações de maneira completamente neutra ou exclusivamente racional. Experiências anteriores, valores, crenças, vínculos afetivos e pertencimentos grupais influenciam continuamente a interpretação dos acontecimentos. A literatura sobre pós-verdade mostra que conteúdos capazes de confirmar convicções previamente estabelecidas tendem a encontrar menor resistência cognitiva, enquanto informações incompatíveis com identidades sociais frequentemente são submetidas a níveis mais elevados de questionamento. As plataformas digitais potencializam essas tendências ao oferecer ambientes altamente personalizados, nos quais diferentes usuários recebem conjuntos distintos de informações conforme seus históricos de interação (Mcintyre, 2018; Pariser, 2012).

Nesse contexto, adquire especial importância a distinção entre bolhas informacionais (*filter bubbles*) e câmaras de eco (*echo chambers*), conceitos frequentemente utilizados como sinônimos, embora designem processos parcialmente distintos. As bolhas informacionais referem-se principalmente aos efeitos produzidos pela personalização algorítmica, mediante a qual sistemas computacionais selecionam conteúdos considerados compatíveis com os interesses previamente identificados de cada usuário. As câmaras de eco, por sua vez, dizem respeito às dinâmicas sociais de interação entre indivíduos que compartilham visões semelhantes de mundo e reforçam mutuamente suas interpretações da realidade. Enquanto a bolha enfatiza a seleção tecnológica da informação, a câmara de eco evidencia os processos sociais de reforço identitário e validação recíproca das crenças. *Na prática, ambos os fenômenos frequentemente coexistem e se fortalecem mutuamente* (Pariser, 2012; Sunstein, 2018).

Contudo, parte da literatura recente adverte que essas categorias não devem ser utilizadas de maneira excessivamente determinista. Estudos empíricos demonstram que muitos usuários continuam expostos a conteúdos divergentes mesmo em ambientes altamente personalizados. A circulação de informações nas plataformas permanece significativamente mais diversa do que frequentemente sugerem interpretações simplificadas sobre isolamento informacional absoluto. O problema central não reside necessariamente na completa ausência de perspectivas diferentes, mas na forma desigual como essas perspectivas são apresentadas, interpretadas e incorporadas pelos sujeitos. A simples exposição

a informações divergentes não garante abertura ao diálogo, podendo inclusive intensificar processos de polarização quando acompanhada por contextos de elevada hostilidade política ou forte identificação grupal (Flaxman; Goel; Rao, 2016; Cesarino, 2021).

A polarização constitui, portanto, elemento fundamental para compreender os efeitos contemporâneos das plataformas digitais. Em sociedades altamente polarizadas, informações deixam progressivamente de ser avaliadas apenas segundo critérios de consistência empírica ou qualidade argumentativa e passam a funcionar como marcadores de identidade coletiva. Compartilhar determinado conteúdo pode significar, simultaneamente, expressar pertencimento político, lealdade ideológica, posicionamento moral ou reconhecimento social dentro de um grupo específico. Nessas circunstâncias, a circulação da informação adquire função simbólica que ultrapassa significativamente seu conteúdo factual. O compartilhamento torna-se ato de identificação coletiva tanto quanto de comunicação (Cesarino, 2021; Recuero, 2018).

Essa característica aproxima diretamente a discussão sobre redes sociais da problemática da pós-verdade desenvolvida na seção anterior. A eficácia de muitos conteúdos não decorre exclusivamente de sua correspondência com os fatos, mas de sua capacidade de produzir identificação emocional, reforçar vínculos comunitários e oferecer interpretações coerentes com determinadas narrativas coletivas. As plataformas algorítmicas ampliam essa dinâmica ao favorecer a rápida circulação de conteúdos capazes de mobilizar emoções intensas, especialmente indignação, medo, humor, surpresa ou pertencimento grupal. O sucesso comunicacional passa a depender menos da precisão informacional e mais da capacidade de gerar elevados níveis de engajamento, compartilhamento e reação emocional (Mcintyre, 2018; Van Dijck, 2013).

A literatura sobre comunicação digital demonstra que conteúdos emocionalmente intensos apresentam maior probabilidade de circulação nas plataformas. Esse fenômeno não significa que os algoritmos sejam programados especificamente para privilegiar informações falsas. O que ocorre é que muitos conteúdos desinformativos são elaborados justamente para explorar características psicológicas que favorecem maior engajamento. Títulos alarmistas, narrativas conspiratórias, simplificações extremas, imagens impactantes e dis-

cursos moralizantes frequentemente despertam respostas emocionais rápidas, aumentando sua capacidade de compartilhamento. A arquitetura orientada pelo engajamento não cria essas estratégias discursivas, mas pode amplificar significativamente sua difusão quando elas produzem elevados índices de interação (Gillespie, 2018; Zuboff, 2019).

Outro aspecto decisivo refere-se à velocidade de circulação da informação. Nas plataformas digitais, conteúdos podem alcançar milhões de pessoas em intervalos extremamente reduzidos, frequentemente antes que processos de verificação, contextualização ou contestação consigam acompanhar sua propagação. A rapidez constitui característica estrutural do ambiente digital contemporâneo. Em muitos casos, a correção posterior de informações falsas alcança públicos significativamente menores do que aqueles expostos à mensagem original. Essa assimetria temporal favorece a consolidação inicial de determinadas narrativas, cuja permanência social pode independender de posterior desmentido factual (Vossoughi; Roy; Aral, 2018).

Ao mesmo tempo, a lógica da conectividade permanente modifica profundamente a experiência cotidiana dos sujeitos. As plataformas deixam de funcionar apenas como ferramentas utilizadas em momentos específicos do dia para integrar continuamente processos de trabalho, educação, entretenimento, relações afetivas, consumo, participação política e acesso à informação. Essa integração produz ambientes comunicacionais caracterizados pela simultaneidade de múltiplos fluxos informacionais, nos quais notícias, opiniões, publicidade, conteúdos científicos, entretenimento, experiências pessoais e discursos políticos aparecem frequentemente misturados na mesma interface. A diferenciação entre fontes especializadas e manifestações individuais torna-se menos evidente para o usuário, ampliando a complexidade dos processos de avaliação da credibilidade das informações (Castells, 2009; Recuero, 2018).

É importante reconhecer, contudo, que a influência das plataformas não elimina a capacidade crítica dos sujeitos. Usuários continuam selecionando conteúdos, estabelecendo vínculos, desenvolvendo estratégias de verificação e reinterpretando mensagens conforme suas experiências pessoais. A literatura contemporânea rejeita tanto interpretações deterministas, que atribuem às tecnologias controle absoluto sobre o comportamento humano, quanto perspectivas

excessivamente voluntaristas, que ignoram os efeitos estruturais da arquitetura digital. As plataformas condicionam possibilidades de ação, mas não determinam integralmente as práticas comunicacionais. Seus efeitos emergem da interação permanente entre arquitetura tecnológica, interesses econômicos, disposições cognitivas, contextos culturais e práticas sociais (Boyd, 2011; Kitchin, 2017).

Essa compreensão permite superar explicações simplificadas que atribuem a disseminação da pós-verdade exclusivamente às redes sociais. As plataformas digitais constituem condição importante para sua amplificação, mas não sua origem exclusiva. Processos históricos de polarização política, crise de confiança nas instituições, transformações econômicas, disputas identitárias e mudanças nas formas de produção do conhecimento antecederam a expansão das tecnologias digitais e continuam influenciando significativamente seus efeitos. As plataformas potencializam essas dinâmicas ao reorganizar a infraestrutura comunicacional da sociedade, tornando mais rápida, personalizada, permanente e economicamente orientada a circulação das informações (Castells, 2009; Cesarino, 2021).

Sob essa perspectiva, a principal contribuição das redes sociais contemporâneas para a compreensão da pós-verdade não consiste simplesmente na existência de ambientes digitais de comunicação, mas na transformação dessas redes em plataformas algorítmicas orientadas pela dataficação, pela economia da atenção e pela modulação comportamental. A articulação entre conectividade global, processamento massivo de dados, personalização algorítmica e modelos econômicos baseados na monetização da atenção cria um ambiente particularmente favorável à circulação acelerada de narrativas concorrentes, emocionalmente mobilizadoras e frequentemente desvinculadas de procedimentos tradicionais de verificação factual. A arquitetura sociotécnica das plataformas passa, assim, a integrar as condições estruturais que sustentam a reorganização contemporânea dos regimes de verdade.

Essa transformação adquire relevância ainda maior diante da emergência das inteligências artificiais generativas. Se as plataformas digitais reorganizaram profundamente a circulação da informação, as novas IAs ampliam significativamente a capacidade de produzir textos, imagens, vídeos, áudios e interações sintéticas em escala praticamente ilimitada. A combinação entre infraestruturas algorítmicas de distribuição e sistemas capazes de automatizar a produção de

conteúdos inaugura uma nova etapa da ecologia informacional contemporânea. As redes sociais deixam de atuar apenas como ambientes de circulação e passam a integrar ecossistemas nos quais produção, personalização, recomendação e difusão de conteúdos podem ser amplamente automatizadas. É justamente essa convergência entre plataformas digitais e inteligências artificiais generativas que constitui o núcleo analítico da presente pesquisa e que será objeto da próxima seção deste estudo.

A reconstrução histórica realizada ao longo deste ensaio evidencia que o conceito de redes sociais percorreu uma trajetória intelectual significativamente mais ampla do que aquela frequentemente presente no debate público contemporâneo. Desde as primeiras formulações sociológicas de Georg Simmel acerca das formas de interação, passando pela sociometria de Jacob Levy Moreno, pelos estudos antropológicos de John Barnes, Elizabeth Bott e John Clyde Mitchell, pela consolidação da Análise de Redes Sociais com Barry Wellman, Wasserman e Faust, até a teoria da força dos laços fracos de Mark Granovetter, observa-se um deslocamento progressivo do olhar científico em direção às relações que estruturam a vida coletiva. Muito antes da internet, as redes sociais já constituíam uma categoria analítica capaz de explicar processos de cooperação, circulação de informações, distribuição de recursos, inovação e organização social. A emergência das tecnologias digitais não inaugura essas relações; apenas reorganiza profundamente sua escala, velocidade, permanência e capacidade de articulação global.

Essa reorganização alcança novo patamar com a formulação da sociedade em rede proposta por Manuel Castells. Ao demonstrar que a lógica reticular passa a constituir a principal morfologia das sociedades informacionais, Castells desloca a discussão das redes sociais para um plano estrutural, no qual economia, política, ciência, cultura e comunicação passam a operar crescentemente por meio de conexões distribuídas. A internet, nessa perspectiva, não representa uma ruptura absoluta com a história das relações sociais, mas a infraestrutura tecnológica que potencializa uma lógica organizacional já presente em diferentes momentos da experiência humana. O advento da Web 2.0 e dos *Social Network Sites* amplia essa transformação ao tornar visíveis, persistentes e navegáveis relações que anteriormente permaneciam restritas a contextos locais ou institucionais.

Entretanto, a principal inflexão conceitual identificada neste capítulo ocorre quando os antigos sites de redes sociais deixam de constituir apenas espaços de interação interpessoal para converterem-se em plataformas digitais. A partir das contribuições de José van Dijck, Tarleton Gillespie, Nick Srnicek e Shoshana Zuboff, torna-se possível compreender que essas infraestruturas já não operam exclusivamente como ambientes destinados ao compartilhamento de conteúdos, mas como sistemas sociotécnicos responsáveis pela coleta massiva de dados, pela personalização algorítmica da informação, pela modulação da atenção e pela organização de fluxos comunicacionais em escala global. As plataformas passam a exercer funções anteriormente distribuídas entre diferentes instituições sociais, reorganizando os processos de produção, circulação, validação e consumo da informação.

Essa mudança altera também a natureza das próprias relações sociais. Se, inicialmente, as redes eram compreendidas como estruturas constituídas pelas conexões estabelecidas entre indivíduos, atualmente essas conexões desenvolvem-se em ambientes cuja arquitetura tecnológica participa ativamente de sua organização. Algoritmos, métricas de engajamento, sistemas de recomendação, mecanismos de personalização e modelos econômicos baseados na monetização da atenção deixam de constituir elementos periféricos para integrar a própria dinâmica relacional. A interação social permanece sendo produzida pelos sujeitos, porém ocorre em ecossistemas digitais que influenciam continuamente sua visibilidade, intensidade, permanência e capacidade de difusão.

Sob essa perspectiva, torna-se insuficiente compreender as redes sociais apenas como instrumentos tecnológicos ou espaços neutros de comunicação. Elas constituem hoje uma infraestrutura informacional capaz de reorganizar significativamente os modos pelos quais indivíduos, grupos e instituições produzem conhecimento, constroem identidades, disputam legitimidade e atribuem significado aos acontecimentos sociais. A arquitetura das plataformas não determina integralmente os comportamentos humanos, mas estabelece condições estruturais que favorecem determinadas formas de circulação da informação em detrimento de outras. Nesse sentido, a compreensão contemporânea das redes sociais exige integrar simultaneamente dimensões sociológicas, comunicacionais, econômicas, computacionais e políticas.

É precisamente essa evolução histórica que permite compreender por que as redes sociais ocupam posição central na problemática da pós-verdade. A intensificação da conectividade global, associada à personalização algorítmica, à economia da atenção, à dataficação e à modulação comportamental, produz um ambiente particularmente propício à rápida circulação de narrativas concorrentes, à fragmentação dos referenciais comuns de validação da informação e à crescente influência de conteúdos organizados prioritariamente em função de seu potencial de engajamento. A transformação das redes sociais em plataformas algorítmicas não constitui, portanto, a origem exclusiva da pós-verdade, mas representa uma das condições estruturais que ampliam significativamente sua capacidade de difusão e de permanência na esfera pública contemporânea.

Essa conclusão conduz diretamente ao próximo movimento analítico desta pesquisa. Se as plataformas digitais reorganizaram profundamente os processos de circulação, seleção e personalização da informação, a emergência das inteligências artificiais generativas inaugura uma etapa qualitativamente distinta dessa transformação. Pela primeira vez, a mesma infraestrutura capaz de distribuir conteúdos em escala planetária passa a articular-se com sistemas igualmente capazes de produzi-los de maneira automatizada, personalizada e praticamente ilimitada. Compreender essa convergência entre plataformas digitais e inteligências artificiais generativas torna-se, portanto, condição indispensável para analisar os novos efeitos da pós-verdade no século XXI, objetivo que orientará a discussão desenvolvida na seção seguinte.

REFERÊNCIAS

BARNES, John A. Class and committees in a Norwegian island parish. **Human Relations**, London, v. 7, n. 1, p. 39–58, 1954. DOI: <https://doi.org/10.1177/001872675400700102>. Disponível em: <https://doi.org/10.1177/001872675400700102>. Acesso em: 14 jul. 2026.

BOURDIEU, Pierre. Le capital social: notes provisoires. **Actes de la Recherche en Sciences Sociales**, Paris, n. 31, p. 2–3, 1980. DOI: <https://doi.org/10.3406/arss.1980.2069>. Disponível em: <https://doi.org/10.3406/arss.1980.2069>. Acesso em: 14 jul. 2026.

BOTT, Elizabeth. *Family and social network*. London: Tavistock Publications, 1957.

BOYD, Danah. Social network sites as networked publics: affordances, dynamics, and implications. In: PAPACHARISSI, Zizi (org.). **A networked self: identity, community, and culture on social network sites**. New York: Routledge, 2011. p. 39–58.

BOYD, Danah; ELLISON, Nicole B. Social network sites: definition, history, and scholarship. **Journal of computer-mediated communication**, Oxford, v. 13, n. 1, p. 210–230, 2007. DOI: <https://doi.org/10.1111/j.1083-6101.2007.00393.x>. Disponível em: <https://doi.org/10.1111/j.1083-6101.2007.00393.x>. Acesso em: 14 jul. 2026.

CASTELLS, Manuel. **A sociedade em rede**. 6. ed. São Paulo: Paz e Terra, 1999.

CASTELLS, Manuel. **A galáxia da internet: reflexões sobre a internet, os negócios e a sociedade**. Rio de Janeiro: Jorge Zahar, 2003.

CASTELLS, Manuel. **Communication power**. Oxford: Oxford University Press, 2009.

CESARINO, Letícia. Pós-verdade e a crise do sistema de peritos: uma explicação cibernética. **Ilha: Revista de Antropologia**, Florianópolis, v. 23, n. 1, p. 73-96, 2021.

COLEMAN, James S. Social capital in the creation of human capital. **American Journal of Sociology**, Chicago, v. 94, p. S95-S120, 1988. DOI: <https://doi.org/10.1086/228943>.

DAVENPORT, Thomas H.; BECK, John C. **The attention economy: understanding the new currency of business**. Boston: Harvard Business School Press, 2001.

FLAXMAN, Seth; GOEL, Sharad; RAO, Justin. Filter bubbles, echo chambers, and online news consumption. **Public Opinion Quarterly**, Oxford, v. 80, p. 298-320, 2016. DOI: <https://doi.org/10.1093/poq/nfw006>.

GILLESPIE, Tarleton. **Custodians of the internet: platforms, content moderation, and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018.

GRANOVETTER, Mark S. The strength of weak ties. **American Journal of Sociology**, Chicago, v. 78, n. 6, p. 1360-1380, 1973. DOI: <https://doi.org/10.1086/225469>.

KAPLAN, Andreas M.; HAENLEIN, Michael. Users of the world, unite! The challenges and opportunities of social media. **Business Horizons, Bloomington**, v. 53, n. 1, p. 59-68, 2010. DOI: <https://doi.org/10.1016/j.bushor.2009.09.003>.

KITCHIN, Rob. Thinking critically about and researching algorithms. **Information, Communication & Society**, London, v. 20, n. 1, p. 14-29, 2017. DOI: <https://doi.org/10.1080/1369118X.2016.1154087>.

LÉVY, Pierre. **Cibercultura**. São Paulo: Editora 34, 1999.

McINTYRE, Lee. **Post-truth**. Cambridge, MA: MIT Press, 2018.

MCPHERSON, Miller; SMITH-LOVIN, Lynn; COOK, James M. Birds of a feather: homophily in social networks. **Annual Review of Sociology**, Palo Alto, v. 27, p. 415-444, 2001. DOI: <https://doi.org/10.1146/annurev.soc.27.1.415>.

MITCHELL, J. Clyde (org.). **Social networks in urban situations: analyses of personal relationships in central african towns**. Manchester: Manchester University Press, 1969.

MORENO, Jacob Levy. **Who shall survive? A new approach to the problem of human interrelations**. Washington, DC: Nervous and Mental Disease Publishing Company, 1934.

O'REILLY, Tim. **What is Web 2.0: design patterns and business models for the next generation of software**. Sebastopol: O'Reilly Media, 2005.

PARISER, Eli. **The filter bubble: what the Internet is hiding from you**. New York: Penguin Press, 2011.

PASQUALE, Frank. **The black box society: the secret algorithms that control money and information**. Cambridge, MA: Harvard University Press, 2015.

PUTNAM, Robert D. **Bowling alone: the collapse and revival of American community**. New York: Simon & Schuster, 2000.

RECUERO, Raquel. **Redes sociais na internet**. 2. ed. Porto Alegre: Sulina, 2018.

SIMMEL, Georg. **Questões fundamentais da sociologia: indivíduo e sociedade**. Rio de Janeiro: Jorge Zahar, 2006.

SIMON, Herbert A. Designing organizations for an information-rich world. In: GREENBERGER, Martin (org.). **Computers, communications, and the public interest**. Baltimore: Johns Hopkins University Press, 1971. p. 37-52.

SRNICEK, Nick. **Platform capitalism**. Cambridge: Polity Press, 2017.

SUNSTEIN, Cass R. **#Republic: divided democracy in the age of social media**. Princeton: Princeton University Press, 2018.

VAN DIJCK, José. **The culture of connectivity: a critical history of social media**. Oxford: Oxford University Press, 2013. DOI: <https://doi.org/10.1093/acprof:oso/9780199970773.001.0001>. Disponível em: <https://doi.org/10.1093/acprof:oso/9780199970773.001.0001>. Acesso em: 14 jul. 2026.

VAN DIJCK, José. Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. **Surveillance & Society**, Kingston, v. 12, n. 2, p. 197–208, 2014. DOI: <https://doi.org/10.24908/ss.v12i2.4776>. Disponível em: <https://doi.org/10.24908/ss.v12i2.4776>. Acesso em: 14 jul. 2026.

VAN DIJCK, José; POELL, Thomas. Understanding social media logic. **Media and Communication**, Lisboa, v. 1, n. 1, p. 2–14, 2013. DOI: <https://doi.org/10.17645/mac.v1i1.70>. Disponível em: <https://doi.org/10.17645/mac.v1i1.70>. Acesso em: 14 jul. 2026.

VAN DIJCK, José; POELL, Thomas; DE WAAL, Martijn. **The platform society: public values in a connective world**. Oxford: Oxford University Press, 2018.

VOSOUGHI, Soroush; ROY, Deb; ARAL, Sinan. The spread of true and false news online. **Science**, Washington, DC, v. 359, n. 6380, p. 1146–1151, 2018. DOI: <https://doi.org/10.1126/science.aap9559>. Disponível em: <https://doi.org/10.1126/science.aap9559>. Acesso em: 14 jul. 2026.

WASSERMAN, Stanley; FAUST, Katherine. **Social network analysis: methods and applications**. Cambridge: Cambridge University Press, 1994.

WELLMAN, Barry. Structural analysis: from method and metaphor to theory and substance. In: WELLMAN, Barry; BERKOWITZ, S. D. (org.). **Social structures: a network approach**. Cambridge: Cambridge University Press, 1988. p. 19–61.

ZUBOFF, Shoshana. **The age of surveillance capitalism: the fight for a human future at the new frontier of power**. New York: PublicAffairs, 2019.

CAPÍTULO 3

MÁQUINAS QUE PRODUZEM SIGNOS: UMA INVESTIGAÇÃO SOBRE A ORIGEM, O DESENVOLVIMENTO E OS EFEITOS CONTEMPORÂNEOS DAS INTELIGÊNCIAS ARTIFICIAIS GENERATIVAS

A expressão “inteligência artificial generativa” adquiriu ampla circulação pública a partir de 2022, especialmente após a disponibilização de sistemas capazes de produzir textos, imagens, códigos computacionais, músicas, vozes e vídeos mediante instruções formuladas em linguagem cotidiana. A velocidade dessa popularização favoreceu a impressão de que se trataria de uma tecnologia inteiramente nova, surgida de maneira quase instantânea com ferramentas como *ChatGPT*, *Midjourney*, *Stable Diffusion* ou *DALL-E*. Essa interpretação, entretanto, obscurece uma história científica muito mais extensa. As atuais inteligências artificiais generativas resultam da convergência entre décadas de pesquisas em inteligência artificial, estatística, aprendizagem de máquina, redes neurais artificiais, processamento de linguagem natural, visão computacional, teoria da informação e desenvolvimento de infraestruturas computacionais capazes de processar enormes conjuntos de dados. O que se tornou novo no início da década de 2020 não foi a ideia de fazer máquinas produzirem representações, mas a escala, a qualidade, a acessibilidade e a multiplicidade de modalidades alcançadas por esses sistemas (Russell; Norvig, 2021; Goodfellow; Bengio; Courville, 2016; UNESCO, 2024).

Em sentido amplo, a inteligência artificial generativa pode ser compreendida como uma categoria de sistemas computacionais projetados para produzir novos conteúdos a partir dos padrões identificados nos dados utilizados em seu treinamento. Esses conteúdos podem assumir a forma de textos, imagens, sons, vídeos, modelos tridimensionais, sequências biológicas, códigos de programação e combinações multimodais. O termo “novo”, contudo, precisa ser empregado com precisão. Os sistemas não criam a partir de uma experiência consciente, de uma intenção subjetiva ou de uma compreensão humana do mundo. Eles geram configurações estatisticamente plausíveis com base nas

regularidades aprendidas durante o treinamento. A produção não consiste em copiar necessariamente um único exemplo previamente armazenado, embora memorizações e reproduções possam ocorrer, mas em calcular e recombinaar padrões de modo a formar uma saída compatível com a distribuição probabilística modelada pelo sistema (Goodfellow; Bengio; Courville, 2016; OECD, 2024; UNESCO, 2024).

Essa definição distingue a IA generativa de sistemas predominantemente classificatórios, preditivos ou discriminativos. Um sistema discriminativo pode, por exemplo, receber uma imagem e estimar se ela representa um gato ou um cachorro; um modelo generativo busca aprender características da distribuição dos dados de modo que possa produzir uma imagem nova que se assemelhe, estatisticamente, às imagens de gatos observadas durante o treinamento. Da mesma forma, um sistema tradicional de análise textual pode classificar uma mensagem como positiva ou negativa, enquanto um modelo generativo pode produzir uma nova mensagem adequada a determinado contexto, gênero textual ou estilo linguístico. Essa distinção não estabelece uma separação absoluta, pois sistemas contemporâneos combinam funções de classificação, previsão, recomendação e geração, mas ajuda a compreender por que a IA generativa representa uma transformação relevante: ela desloca a máquina da posição exclusiva de analisadora de conteúdos para a de produtora de artefatos simbólicos aparentemente inéditos (Goodfellow; Bengio; Courville, 2016; OECD, 2024).

O fundamento matemático das tecnologias generativas contemporâneas encontra-se nos chamados modelos generativos, cuja finalidade consiste em aprender propriedades da distribuição probabilística dos dados. Esses modelos procuram representar relações, padrões e dependências presentes em um conjunto de exemplos, de modo que possam posteriormente produzir novas amostras compatíveis com aquilo que aprenderam. Em um modelo de linguagem, por exemplo, a tarefa fundamental consiste em estimar a probabilidade de uma palavra, parte de palavra ou unidade textual ocorrer em determinado contexto. Nos modelos autorregressivos, o texto é produzido sequencialmente: cada novo elemento é calculado a partir dos elementos anteriores e das probabilidades construídas durante o treinamento. Uma resposta extensa pode, assim, emergir de uma sucessão de previsões locais, sem que o sistema possua necessaria-

mente uma representação consciente da totalidade do argumento que apresenta (Brown et al., 2020; OpenAI, 2023).

A compreensão dessa natureza probabilística é indispensável para evitar dois equívocos opostos. O primeiro consiste em imaginar que a IA generativa funciona apenas como um banco de dados que recupera frases prontas. Embora esses sistemas possam memorizar conteúdos e reproduzir trechos presentes em seus dados, sua operação habitual envolve a geração dinâmica de sequências a partir de padrões estatísticos. O segundo equívoco consiste em atribuir-lhes uma inteligência equivalente à consciência humana, como se a fluidez linguística demonstrasse compreensão plena, intenção, julgamento moral ou compromisso espontâneo com a verdade. A capacidade de produzir um texto coerente não equivale necessariamente à capacidade de compreender ontologicamente aquilo que está sendo afirmado. O modelo pode organizar com notável competência as regularidades formais de uma linguagem e, simultaneamente, produzir informações factualmente incorretas, referências inexistentes ou explicações incompatíveis com a realidade (Bender et al., 2021; Floridi; Chiriatti, 2020).

O caráter generativo da inteligência artificial, portanto, não começa com os grandes modelos de linguagem. Desde as primeiras fases da computação, pesquisadores buscaram desenvolver programas capazes de simular conversações, compor músicas, produzir desenhos ou criar sequências simbólicas. O programa ELIZA, desenvolvido por Joseph Weizenbaum (1966) na década de 1960, tornou-se um dos exemplos mais conhecidos. Utilizando regras relativamente simples de reconhecimento de padrões e reformulação de frases, o sistema simulava uma conversa terapêutica e produzia nos usuários uma forte impressão de compreensão. Embora tecnicamente distante dos modelos atuais, ELIZA revelou um fenômeno que permanece central: seres humanos tendem a atribuir intencionalidade, sensibilidade e compreensão a máquinas capazes de responder em linguagem aparentemente significativa. O chamado “efeito ELIZA” demonstra que a força social de um sistema não depende apenas de suas capacidades técnicas, mas também das interpretações e projeções humanas dirigidas a ele (Turkle, 2011).

O projeto histórico da inteligência artificial oscilou, durante décadas, entre abordagens simbólicas e conexionistas. A inteligência artificial simbólica

buscava representar conhecimentos mediante regras explícitas, estruturas lógicas e sistemas formais manipuláveis pelo computador. Nessa tradição, produzir uma resposta adequada exigia que programadores descrevessem previamente as relações relevantes entre conceitos e situações. Já as abordagens conexionistas inspiraram-se, de maneira abstrata, na organização de redes neurais, procurando fazer com que sistemas aprendessem padrões a partir de exemplos. Nenhuma dessas vertentes permaneceu absolutamente isolada, mas a expansão contemporânea da IA generativa está fortemente associada ao amadurecimento das técnicas de aprendizagem profunda, ou *deep learning*, que tornaram possível treinar redes neurais com muitas camadas sobre quantidades massivas de dados (Goodfellow; Bengio; Courville, 2016; Russell; Norvig, 2021).

As redes neurais artificiais são modelos computacionais constituídos por unidades matemáticas interligadas cujos parâmetros são ajustados durante o treinamento. De maneira simplificada, o sistema recebe exemplos, produz resultados, compara-os com os resultados esperados e altera seus parâmetros para reduzir o erro. Repetido em grande escala, esse processo permite que o modelo reconheça regularidades extremamente complexas sem que todas as regras precisem ser programadas explicitamente. O desenvolvimento de unidades de processamento gráfico, a disponibilidade de bases de dados extensas, o aprimoramento de algoritmos de treinamento e a expansão da capacidade de armazenamento tornaram viável o treinamento de modelos progressivamente maiores, favorecendo avanços expressivos em reconhecimento de fala, tradução automática, visão computacional e processamento de linguagem natural (Lecun; Bengio; Hinton, 2015; Goodfellow; Bengio; Courville, 2016).

Um marco decisivo na história recente da geração computacional ocorreu em 2014, quando Ian Goodfellow e colaboradores apresentaram as redes generativas adversariais, conhecidas pela sigla GAN, de *Generative Adversarial Networks*. Essa arquitetura organiza o treinamento como uma disputa entre duas redes neurais: um gerador procura produzir amostras semelhantes aos dados reais, enquanto um discriminador procura distinguir as amostras verdadeiras daquelas fabricadas pelo gerador. À medida que ambos competem, o gerador aperfeiçoa progressivamente sua capacidade de criar resultados convincentes. O modelo foi originalmente demonstrado na geração de imagens, mas sua

lógica influenciou uma ampla variedade de aplicações posteriores, incluindo síntese visual, transformação de estilos e criação de rostos humanos artificiais (Goodfellow et al., 2014).

As GANs contribuíram decisivamente para o desenvolvimento de imagens sintéticas cada vez mais realistas e, conseqüentemente, para a expansão das preocupações relacionadas aos *deepfakes*. O termo passou a designar conteúdos audiovisuais manipulados ou integralmente sintetizados por sistemas de aprendizagem profunda, capazes de substituir rostos, reproduzir vozes, simular movimentos e atribuir falas ou ações a pessoas reais. A possibilidade de produzir evidências audiovisuais convincentes sem acontecimento correspondente representa uma ruptura importante na relação social com o documento. Fotografias, áudios e vídeos jamais foram provas absolutamente neutras, mas tradicionalmente possuíam elevado valor testemunhal. A geração sintética reduz os custos e as competências necessárias à falsificação, ampliando tanto a capacidade de fabricar acontecimentos quanto a possibilidade de negar registros autênticos sob a alegação de que seriam artificialmente produzidos (Chesney; Citron, 2019; Westerlund, 2019).

Paralelamente ao desenvolvimento das GANs, outra trajetória técnica transformava profundamente o processamento de linguagem. Modelos de linguagem probabilísticos já eram utilizados para prever sequências textuais, mas enfrentavam dificuldades para relacionar elementos distantes em um texto e para processar grandes quantidades de informação com eficiência. Redes recorrentes e arquiteturas do tipo LSTM representaram avanços importantes, pois permitiam conservar informações ao longo de uma sequência. Contudo, seu processamento predominantemente sequencial dificultava a paralelização em larga escala e limitava o treinamento de modelos ainda maiores. A superação parcial dessas restrições ocorreria com o desenvolvimento da arquitetura Transformer (Hochreiter; Schmidhuber, 1997; Vaswani et al., 2017).

Apresentada por Ashish Vaswani e colaboradores no artigo “*Attention Is All You Need*”, de 2017, a arquitetura Transformer substituiu grande parte da recorrência por mecanismos de atenção. Esses mecanismos permitem que o modelo atribua diferentes pesos às relações entre os elementos de uma sequência, identificando quais partes do contexto são mais relevantes para interpretar

ou produzir determinado elemento. Além de melhorar o tratamento de dependências textuais distantes, a arquitetura tornou o treinamento mais paralelizável e eficiente. Embora tenha sido inicialmente desenvolvida para tradução automática, tornou-se a base de grande parte dos modelos contemporâneos de linguagem e, posteriormente, de sistemas multimodais capazes de combinar texto, imagem, áudio e vídeo (Vaswani et al., 2017).

A importância do Transformer não decorre apenas de uma melhoria de desempenho. Sua arquitetura criou condições técnicas para o desenvolvimento de modelos extensos, pré-treinados em grandes corpora e posteriormente adaptáveis a diferentes tarefas. Em vez de construir um sistema independente para tradução, outro para resumo e outro para resposta a perguntas, tornou-se possível treinar um modelo geral sobre enormes coleções textuais e utilizá-lo como base para diversas aplicações. Essa mudança favoreceu a emergência dos chamados modelos fundacionais, sistemas treinados em dados amplos que podem ser adaptados a múltiplos usos posteriores. A flexibilidade desses modelos transforma a IA generativa em uma tecnologia de caráter geral, capaz de penetrar simultaneamente em diferentes setores econômicos, profissionais, educacionais e culturais (Bommasani et al., 2021).

A sigla GPT, de *Generative Pre-trained Transformer*, expressa três elementos fundamentais dessa mudança. O sistema é “generativo” porque produz novas sequências; é “pré-treinado” porque aprende previamente com grandes conjuntos de dados antes de ser aplicado a tarefas específicas; e utiliza a arquitetura Transformer. O pré-treinamento permite que o modelo desenvolva representações estatísticas da linguagem, reconhecendo regularidades gramaticais, semânticas, discursivas e culturais presentes nos textos. Posteriormente, o modelo pode ser ajustado, instruído ou condicionado para responder perguntas, resumir documentos, traduzir conteúdos, produzir códigos, simular estilos e executar outras tarefas. Essa versatilidade explica por que os grandes modelos de linguagem deixaram rapidamente os laboratórios de pesquisa e passaram a integrar produtos de uso cotidiano.

A publicação do GPT-3, descrito por Brown e colaboradores em 2020, tornou especialmente visível o efeito da ampliação de escala. Com 175 bilhões de parâmetros, o modelo apresentou capacidade de realizar diversas tarefas a

partir de instruções ou de poucos exemplos fornecidos no próprio contexto da interação, sem necessidade de novo treinamento específico para cada atividade. O estudo demonstrou resultados expressivos em tradução, respostas a perguntas, elaboração textual e outras tarefas, mas também reconheceu limitações, vieses e dificuldades metodológicas relacionadas ao treinamento em grandes conjuntos de textos extraídos da internet. Particularmente relevante para a problemática da pós-verdade foi a constatação de que leitores humanos apresentavam dificuldades para distinguir determinadas notícias produzidas pelo modelo de textos escritos por pessoas (Brown et al., 2020).

A expansão das capacidades não decorreu unicamente do aumento no número de parâmetros. O desenvolvimento dos atuais assistentes conversacionais envolveu técnicas adicionais de ajuste supervisionado, aprendizagem a partir de preferências humanas e alinhamento comportamental. Em termos gerais, avaliadores humanos analisam respostas produzidas pelos modelos, comparam alternativas e oferecem sinais que podem ser utilizados para ajustar o sistema em direção a comportamentos considerados mais úteis e seguros. Esse processo tornou as interfaces mais capazes de seguir instruções, sustentar diálogos e recusar certos pedidos. Contudo, alinhamento não significa eliminação de erros, neutralidade ou concordância universal sobre valores. As preferências incorporadas refletem decisões institucionais, critérios técnicos, normas culturais e limites impostos pelos dados e pelos próprios avaliadores humanos (Ouyang et al., 2022).

A disponibilização pública do *ChatGPT*, em novembro de 2022, representou uma mudança social comparável, em alcance, a alguns dos principais momentos da expansão da internet. A inovação não estava apenas no modelo subjacente, mas na interface conversacional simples, capaz de transformar operações computacionais complexas em uma interação acessível por meio da linguagem comum. Milhões de pessoas sem formação em programação passaram a solicitar textos, resumos, explicações, traduções, códigos, planejamentos e análises mediante perguntas formuladas de modo cotidiano. A IA generativa deixava de ser um recurso predominantemente especializado e transformava-se em tecnologia cultural de uso geral, incorporada a atividades profissionais, educacionais, artísticas, administrativas e comunicacionais (UNESCO, 2024; Lima; Gomes; Silva, 2024).

A expressão “novas IAs generativas” deve, portanto, ser compreendida menos como referência a uma única tecnologia e mais como designação de um ecossistema heterogêneo de modelos, infraestruturas e aplicações. Nele coexistem grandes modelos de linguagem, geradores de imagens por difusão, sintetizadores de voz, modelos de vídeo, sistemas multimodais, ferramentas de programação, modelos científicos e agentes capazes de executar sequências de ações. O adjetivo “novas” diz respeito sobretudo à combinação entre escala de treinamento, acesso por linguagem natural, capacidade multimodal, integração às plataformas digitais e possibilidade de produzir conteúdos em quantidade e velocidade sem precedentes. Essa combinação inaugura uma transformação não apenas tecnológica, mas epistemológica: máquinas passam a participar diretamente da produção cotidiana dos signos pelos quais a sociedade descreve, interpreta e disputa a realidade.

A expansão contemporânea das inteligências artificiais generativas não pode ser compreendida apenas pelo aumento da capacidade computacional ou pela evolução dos algoritmos. O que distingue os sistemas atuais de experiências anteriores é a convergência entre diferentes transformações técnicas ocorridas ao longo das duas primeiras décadas do século XXI: a disponibilidade de bases de dados em escala planetária, o desenvolvimento de arquiteturas neurais mais eficientes, o crescimento exponencial do poder computacional, a utilização de unidades especializadas de processamento gráfico, o aperfeiçoamento dos mecanismos de treinamento distribuído e a consolidação de infraestruturas globais de computação em nuvem. Essas condições permitiram o surgimento dos chamados modelos fundacionais (*foundation models*), capazes de servir como base para uma ampla variedade de aplicações posteriores. Diferentemente dos sistemas especializados em uma única tarefa, esses modelos aprendem representações gerais da linguagem, das imagens ou de outras modalidades de dados e podem ser posteriormente adaptados para inúmeras finalidades distintas (Bommasani et al., 2021; Russell; Norvig, 2021).

O conceito de modelo fundacional representa uma mudança importante na própria compreensão da inteligência artificial. Durante grande parte da história da área, predominou a construção de sistemas específicos para problemas específicos. Desenvolvia-se um algoritmo para reconhecer rostos, outro para tradu-

zir idiomas, outro para responder perguntas e outro para resumir documentos. A emergência dos grandes modelos alterou esse paradigma ao demonstrar que um único sistema, treinado sobre enormes conjuntos de dados heterogêneos, pode adquirir competências suficientemente gerais para executar diferentes tarefas mediante instruções apropriadas. O conhecimento deixa de estar rigidamente associado a uma única aplicação e passa a constituir um repertório estatístico reutilizável em múltiplos contextos. Essa flexibilidade aproxima a IA generativa das chamadas tecnologias de propósito geral, cujos efeitos ultrapassam setores específicos e reorganizam amplamente práticas econômicas, científicas, educacionais e culturais (Bommasani et al., 2021).

Essa transformação também modifica profundamente a relação entre treinamento e utilização. O treinamento de um grande modelo exige recursos computacionais extraordinários, envolvendo bilhões ou trilhões de palavras, milhões de imagens e enormes quantidades de energia elétrica, armazenamento e processamento distribuído. Entretanto, uma vez treinado, o modelo pode ser disponibilizado para milhões de usuários simultaneamente, produzindo respostas individualizadas em tempo real. A economia da inteligência artificial desloca-se, assim, da produção repetida de sistemas independentes para o desenvolvimento de infraestruturas centrais reutilizadas em múltiplas aplicações. Esse movimento contribui para a concentração tecnológica em torno de poucas organizações capazes de financiar treinamentos de grande escala, produzindo novas assimetrias econômicas e geopolíticas no desenvolvimento da inteligência artificial (OECD, 2024; UNESCO, 2024).

Outro aspecto decisivo refere-se à natureza dos dados utilizados durante o treinamento. Modelos generativos aprendem padrões presentes nos exemplos fornecidos, razão pela qual sua qualidade depende fortemente da diversidade, extensão e representatividade desses conjuntos de dados. Textos científicos, livros, páginas da internet, códigos de programação, artigos jornalísticos, documentos públicos, imagens, gravações sonoras e inúmeros outros materiais podem integrar o processo de aprendizagem. Contudo, esses dados não constituem um retrato neutro da realidade. Eles refletem desigualdades históricas, assimetrias culturais, vieses linguísticos, exclusões sociais e disputas políticas presentes nas sociedades que os produziram. Consequentemente, modelos trei-

nados sobre esses conjuntos tendem a reproduzir, em maior ou menor grau, regularidades estatísticas que incluem tanto conhecimentos socialmente valiosos quanto preconceitos, estereótipos e distorções historicamente sedimentados (Bender et al., 2021; UNESCO, 2021).

Essa constatação desloca o debate sobre inteligência artificial para além das questões estritamente técnicas. O desempenho de um modelo não depende apenas da sofisticação de sua arquitetura computacional, mas também das escolhas realizadas na seleção, filtragem, curadoria e documentação dos dados utilizados durante o treinamento. Perguntas sobre quem produziu esses dados, quais idiomas predominam, quais culturas permanecem sub-representadas, quais grupos sociais aparecem com maior frequência e quais conhecimentos foram excluídos tornam-se centrais para compreender os resultados produzidos pelos sistemas generativos. A inteligência artificial revela-se, assim, profundamente dependente das condições históricas de produção do conhecimento humano que alimenta seus processos de aprendizagem (Bender et al., 2021; Floridi; Chiriatti, 2020).

Essas questões adquirem relevância ainda maior quando se considera a expansão da multimodalidade. Enquanto os primeiros grandes modelos concentravam-se predominantemente na linguagem escrita, as arquiteturas mais recentes passaram a integrar simultaneamente diferentes modalidades de informação. Um único sistema pode receber textos, imagens, diagramas, áudio ou vídeo como entrada e produzir respostas que combinam diferentes formas de representação. Essa integração amplia significativamente as possibilidades de aplicação da inteligência artificial, permitindo interpretar documentos complexos, analisar imagens médicas, gerar ilustrações, descrever fotografias para pessoas com deficiência visual, sintetizar apresentações multimídia e combinar diferentes linguagens em uma única interação. A multimodalidade aproxima os sistemas computacionais das formas pelas quais os seres humanos tradicionalmente integram diferentes canais perceptivos durante seus processos de comunicação (OpenAI, 2023; UNESCO, 2024).

A geração de imagens constitui um exemplo particularmente expressivo dessa transformação. Enquanto as GANs dominaram durante alguns anos a produção de imagens sintéticas, modelos de difusão passaram progressivamen-

te a ocupar posição central nesse campo. Em termos simplificados, esses sistemas aprendem a reconstruir imagens a partir de processos graduais de remoção de ruído estatístico. Durante o treinamento, o modelo observa sucessivas degradações de imagens reais e aprende a inverter esse processo, reconstruindo representações visualmente coerentes. O resultado é uma geração de imagens frequentemente mais estável, detalhada e controlável do que aquela produzida pelas arquiteturas adversariais clássicas. Ferramentas como *Stable Diffusion* e *DALL-E* tornaram-se exemplos amplamente conhecidos dessa abordagem, permitindo que descrições textuais sejam convertidas em imagens sintéticas de elevada qualidade visual (Rombach et al., 2022).

O desenvolvimento desses modelos evidencia que a inteligência artificial generativa não se limita ao domínio da linguagem. O princípio subjacente consiste em aprender distribuições estatísticas suficientemente complexas para produzir novos objetos compatíveis com determinado universo de dados. Seja na geração de textos, imagens, músicas, estruturas moleculares ou proteínas, o desafio fundamental permanece semelhante: construir representações matemáticas capazes de modelar regularidades presentes nos exemplos observados e utilizá-las para produzir novas combinações. Essa generalidade explica por que diferentes campos científicos passaram a incorporar rapidamente modelos generativos em atividades de pesquisa, inovação tecnológica e desenvolvimento industrial (Goodfellow; Bengio; Courville, 2016; Russell; Norvig, 2021).

Entretanto, a impressionante capacidade de produzir conteúdos sintaticamente coerentes e visualmente convincentes suscita uma questão epistemológica decisiva: em que medida esses sistemas podem ser considerados criativos? A resposta exige distinguir cuidadosamente criatividade humana e geração estatística. Sob determinadas perspectivas da Psicologia Cognitiva e dos estudos da criatividade, criar envolve não apenas produzir algo novo, mas atribuir significado, estabelecer intenções, elaborar experiências subjetivas e situar a produção em contextos históricos e culturais compartilhados. Os modelos generativos, por sua vez, operam mediante transformação probabilística de padrões presentes nos dados. Seus resultados podem ser originais em termos combinatórios, surpreendentes para os observadores e altamente úteis em diferentes contextos, mas não decorrem de consciência reflexiva, desejo, imagi-

nação autobiográfica ou experiência fenomenológica do mundo (Boden, 2004; Floridi; Chiriatti, 2020).

Margaret Boden (2004) propõe uma distinção particularmente útil entre criatividade psicológica e criatividade histórica. Uma ideia pode ser psicologicamente nova para determinado indivíduo mesmo quando já exista anteriormente na cultura, enquanto uma criação historicamente inovadora modifica efetivamente o repertório disponível para toda uma comunidade. Sob essa perspectiva, sistemas generativos podem produzir resultados psicologicamente inéditos para os usuários, combinando elementos de maneiras inesperadas. Todavia, permanece objeto de intenso debate filosófico e científico se tais sistemas podem ser considerados agentes criativos em sentido pleno ou se constituem ferramentas capazes de ampliar as possibilidades criativas dos seres humanos. O centro da discussão desloca-se da pergunta “a máquina é criativa?” para questões mais complexas sobre autoria, colaboração, mediação tecnológica e produção coletiva do conhecimento.

Esses debates tornam-se particularmente relevantes no campo da educação, da ciência e da produção intelectual. A facilidade com que modelos generativos elaboram resumos, ensaios, traduções, códigos computacionais e revisões bibliográficas desafia concepções tradicionais de autoria e avaliação acadêmica. Ao mesmo tempo em que oferecem oportunidades inéditas para acessibilidade, personalização da aprendizagem, apoio à escrita e democratização do acesso ao conhecimento, também levantam preocupações relacionadas ao plágio, à dependência cognitiva, à redução da autonomia intelectual e à dificuldade de distinguir produção humana e produção assistida por inteligência artificial. Organismos internacionais, como a UNESCO, têm enfatizado que a incorporação dessas tecnologias aos sistemas educacionais exige desenvolvimento simultâneo de competências críticas, transparência sobre seu uso e fortalecimento da formação ética de estudantes e professores (UNESCO, 2023; UNESCO, 2024).

Essa preocupação decorre do fato de que a inteligência artificial generativa não constitui apenas uma inovação tecnológica, mas uma tecnologia epistemológica. Ao produzir textos, imagens e argumentos aparentemente plausíveis, esses sistemas passam a participar da construção cotidiana do conhecimento social. Diferentemente dos mecanismos tradicionais de busca, que direcionam

o usuário para documentos previamente existentes, os modelos generativos sintetizam respostas inéditas, reorganizando informações provenientes de múltiplas fontes em uma única narrativa. Essa mudança altera significativamente a relação entre sujeito, documento e conhecimento. O usuário deixa progressivamente de consultar diretamente diferentes fontes para dialogar com um sistema capaz de produzir sínteses imediatas, tornando ainda mais importante compreender os critérios, limites e possibilidades envolvidos nesse processo (Floridi; Chiriatti, 2020; UNESCO, 2024).

A crescente naturalidade das interfaces conversacionais intensifica esse fenômeno. Linguagem fluente, respostas contextualizadas, adaptação ao estilo do interlocutor e capacidade de sustentar diálogos prolongados favorecem a construção de uma experiência comunicacional extremamente próxima da interação humana. Como demonstrado desde o efeito ELIZA, essa proximidade pode estimular projeções antropomórficas, levando usuários a atribuírem compreensão, intenção ou autoridade epistêmica a sistemas cujo funcionamento permanece fundamentalmente probabilístico. A distinção entre competência linguística e compreensão conceitual torna-se, portanto, uma questão central para qualquer investigação sobre os impactos sociais das inteligências artificiais generativas.

É justamente essa transformação epistemológica que prepara a discussão seguinte. Se os grandes modelos generativos modificam profundamente as formas de produzir, organizar e apresentar conhecimentos, torna-se necessário investigar como essas capacidades se articulam aos ambientes algorítmicos analisados na seção anterior. A convergência entre plataformas digitais orientadas pela economia da atenção e sistemas capazes de produzir conteúdos sintéticos em escala inaugura um novo cenário para a circulação da informação. Nesse contexto, compreender os usos sociais, econômicos, científicos e políticos das inteligências artificiais generativas torna-se condição indispensável para analisar seus efeitos sobre a construção contemporânea da verdade, tema que será desenvolvido na sequência destas proposições.

A rápida incorporação das inteligências artificiais generativas às diferentes esferas da vida social decorre, em grande medida, de seu caráter transversal. Diferentemente de tecnologias concebidas para resolver problemas espe-

cíficos, os modelos generativos apresentam potencial para serem incorporados simultaneamente à educação, à saúde, ao direito, à engenharia, ao jornalismo, à administração pública, à pesquisa científica, à indústria criativa, ao desenvolvimento de software e aos mais diversos setores produtivos. Essa característica aproxima a inteligência artificial generativa das chamadas tecnologias de propósito geral, isto é, inovações capazes de modificar transversalmente processos econômicos, organizacionais e culturais. A amplitude de suas aplicações faz com que seus impactos extrapolem significativamente o domínio da computação, alcançando praticamente todas as atividades humanas mediadas pela produção, circulação ou transformação de informações (Brynjolfsson; McAfee, 2017; Bommasani et al., 2021).

No campo científico, a IA generativa inaugura possibilidades inéditas para o tratamento de grandes volumes de informação. Pesquisadores podem utilizá-la para sintetizar literatura especializada, identificar padrões em bases documentais extensas, auxiliar na programação computacional, sugerir hipóteses preliminares, estruturar projetos de pesquisa, produzir traduções técnicas e apoiar a redação inicial de manuscritos científicos. Em áreas como Bioinformática, Química Computacional e Biologia Molecular, modelos generativos vêm sendo empregados na proposição de estruturas proteicas, na descoberta de moléculas com potencial terapêutico e na simulação de interações bioquímicas, reduzindo significativamente o tempo necessário para determinadas etapas da investigação científica. Contudo, esses avanços exigem rigorosa supervisão humana, uma vez que a plausibilidade formal das respostas não garante sua validade empírica ou metodológica (Jumper et al., 2021; Nature, 2023).

Na educação, as inteligências artificiais generativas vêm promovendo uma reorganização das práticas de ensino, aprendizagem e avaliação. Sistemas conversacionais podem oferecer explicações individualizadas, adaptar exemplos ao repertório do estudante, sugerir exercícios com diferentes níveis de complexidade, produzir feedback imediato e auxiliar na organização de planos de estudo personalizados. Tais possibilidades aproximam-se dos princípios da aprendizagem adaptativa e da educação personalizada, historicamente discutidos pela Psicologia Educacional e pela Tecnologia Educacional. Entretanto, a literatura especializada enfatiza que esses recursos não substituem os processos

pedagógicos nem o papel docente. Ao contrário, sua utilização exige o fortalecimento da mediação pedagógica, do pensamento crítico, da alfabetização digital e da capacidade de avaliação das informações produzidas pelas próprias inteligências artificiais (UNESCO, 2023; Holmes; Bialik; Fadel, 2019).

A incorporação da IA generativa aos ambientes escolares também desloca profundamente a compreensão tradicional da autoria. Durante décadas, a produção textual foi concebida como expressão relativamente direta da elaboração intelectual do estudante. Com a emergência dos modelos generativos, torna-se possível produzir textos estruturalmente coerentes, semanticamente organizados e linguisticamente sofisticados mediante poucos comandos escritos em linguagem natural. Essa transformação desafia modelos tradicionais de avaliação baseados predominantemente no produto final da escrita. Em resposta, pesquisadores têm defendido a valorização crescente dos processos de elaboração, da argumentação oral, da construção colaborativa do conhecimento e da explicitação crítica das interações estabelecidas entre estudantes e sistemas de inteligência artificial (UNESCO, 2023; Perkins, 2024).

No âmbito profissional, a IA generativa modifica significativamente a natureza de diferentes atividades cognitivas. Em vez de substituir integralmente determinadas profissões, os estudos mais recentes sugerem que esses sistemas tendem a reorganizar tarefas específicas, automatizando atividades repetitivas e ampliando a produtividade em processos intensivos em informação. Advogados utilizam modelos generativos para elaborar versões preliminares de contratos e pareceres; jornalistas recorrem a sistemas automatizados para sintetizar documentos e organizar dados; programadores recebem sugestões de código em tempo real; designers exploram rapidamente diferentes possibilidades visuais; profissionais da saúde utilizam ferramentas de apoio à documentação clínica. Em praticamente todos esses casos, a inteligência artificial funciona como instrumento de ampliação da capacidade humana, exigindo, entretanto, mecanismos robustos de supervisão, validação e responsabilidade profissional (Brynjolfsson; McAfee, 2017; OECD, 2024).

Essa reorganização do trabalho intelectual encontra paralelo em processos históricos anteriores de transformação tecnológica. Assim como a mecanização modificou profundamente o trabalho físico durante a Revolução Industrial, a

inteligência artificial generativa passa a atuar sobre atividades tradicionalmente associadas ao processamento simbólico. Contudo, a analogia deve ser empregada com cautela. A automatização contemporânea não elimina a necessidade de julgamento humano, especialmente em contextos que envolvem decisões éticas, interpretação jurídica, responsabilidade científica, diagnóstico clínico, planejamento educacional ou formulação de políticas públicas. Quanto maior a complexidade social da tarefa, maior tende a ser a importância da intervenção humana na interpretação, contextualização e validação dos resultados produzidos pelos sistemas computacionais (Russell; Norvig, 2021; UNESCO, 2021).

No setor produtivo, as aplicações das inteligências artificiais generativas ultrapassam amplamente a produção textual. Empresas vêm utilizando esses modelos para desenvolvimento de campanhas publicitárias, atendimento automatizado, criação de protótipos industriais, otimização logística, produção audiovisual, simulação de cenários econômicos e desenvolvimento acelerado de software. A capacidade de produzir rapidamente múltiplas alternativas reduz custos de experimentação e amplia significativamente os processos de inovação incremental. Ao mesmo tempo, essa aceleração intensifica disputas relacionadas aos direitos autorais, à propriedade intelectual, à remuneração de criadores humanos e ao uso de obras protegidas durante o treinamento dos modelos (OECD, 2024; WIPO, 2024).

No campo artístico, a IA generativa inaugura uma das discussões mais complexas da contemporaneidade. Ferramentas capazes de produzir pinturas, ilustrações, composições musicais, roteiros cinematográficos, fotografias sintéticas e vídeos hiper-realistas desafiam concepções tradicionais sobre criatividade, autoria e originalidade. Diferentemente das tecnologias digitais anteriores, que frequentemente funcionavam como instrumentos utilizados pelo artista, os sistemas generativos participam ativamente da elaboração do resultado final. Muitos pesquisadores defendem que essa relação deve ser compreendida em termos de coautoria ou cogeração, enquanto outros sustentam que a autoria permanece essencialmente humana, uma vez que os modelos operam exclusivamente a partir de repertórios culturais produzidos por pessoas. O debate permanece aberto e envolve dimensões jurídicas, filosóficas, estéticas e éticas ainda em consolidação (Boden, 2004; Floridi; Chiriatti, 2020).

Paralelamente às oportunidades, cresce a preocupação internacional com os riscos associados ao uso indiscriminado dessas tecnologias. Um dos desafios mais recorrentes refere-se às chamadas “alucinações” (*hallucinations*), expressão utilizada para designar respostas produzidas pelos modelos que apresentam elevada coerência linguística, mas baixo compromisso factual. O sistema pode gerar referências bibliográficas inexistentes, interpretar incorretamente documentos, combinar informações provenientes de fontes distintas como se pertencessem ao mesmo contexto histórico ou afirmar fatos que simplesmente não ocorreram. O problema não decorre de intenção deliberada de enganar, mas da própria lógica probabilística dos modelos, cuja prioridade consiste em produzir sequências linguisticamente plausíveis e não verificar empiricamente a veracidade de cada afirmação (Bender et al., 2021; OpenAI, 2023).

Essa característica possui implicações particularmente relevantes para a produção do conhecimento científico. A facilidade de obtenção de respostas rápidas pode induzir pesquisadores, estudantes e profissionais a reduzirem procedimentos de verificação documental, conferência de fontes e análise crítica da literatura. Organismos internacionais e revistas científicas têm enfatizado que sistemas generativos podem auxiliar significativamente a escrita acadêmica, mas não substituem a responsabilidade intelectual dos autores sobre a precisão das informações apresentadas. Em consequência, diferentes editoras passaram a exigir declaração explícita sobre o uso de inteligência artificial durante a elaboração de manuscritos, reafirmando que a autoria científica permanece inseparável da responsabilidade humana pela integridade do trabalho publicado (Nature, 2023; COPE, 2023).

Outra preocupação crescente refere-se à concentração econômica da inteligência artificial. O treinamento de grandes modelos demanda investimentos financeiros, infraestrutura computacional e acesso a bases de dados que permanecem concentrados em número relativamente reduzido de empresas e instituições. Essa concentração tende a ampliar desigualdades tecnológicas entre países, organizações e comunidades científicas, produzindo novas formas de dependência informacional. Questões relacionadas à soberania digital, transparência algorítmica, governança dos dados e acesso democrático às tecnologias passam a integrar o centro das discussões internacionais sobre inteligência artificial responsável (UNESCO, 2021; OECD, 2024).

A dimensão ética dessas transformações tornou-se suficientemente relevante para motivar a elaboração de documentos internacionais específicos. A Recomendação sobre a Ética da Inteligência Artificial, aprovada pela UNESCO em 2021, propõe princípios voltados à proteção dos direitos humanos, da diversidade cultural, da transparência, da responsabilidade, da inclusão, da sustentabilidade e da supervisão humana dos sistemas inteligentes. O documento parte do reconhecimento de que a inteligência artificial possui elevado potencial para contribuir com o desenvolvimento humano, mas enfatiza que sua utilização deve estar subordinada aos valores democráticos, à dignidade da pessoa humana e à preservação da autonomia individual. A inovação tecnológica, nessa perspectiva, não constitui finalidade em si mesma, mas instrumento cuja legitimidade depende de seus impactos sociais e éticos (UNESCO, 2021).

Essa orientação revela uma mudança importante na forma de compreender a própria inteligência artificial. O debate contemporâneo desloca-se progressivamente da pergunta “o que essas tecnologias conseguem fazer?” para questões mais abrangentes sobre “como elas reorganizam práticas sociais, relações de poder e formas de produção do conhecimento?”. Essa inflexão aproxima os estudos sobre IA generativa das discussões desenvolvidas anteriormente acerca das plataformas digitais. Em ambos os casos, a tecnologia deixa de ser analisada apenas por suas características técnicas e passa a ser compreendida como componente de ecossistemas sociotécnicos complexos, nos quais infraestrutura computacional, interesses econômicos, práticas culturais e decisões políticas encontram-se profundamente articulados.

Sob essa perspectiva, a inteligência artificial generativa representa mais do que um novo conjunto de ferramentas computacionais. Ela inaugura uma nova etapa da mediação tecnológica da cultura, na qual sistemas automatizados passam a participar diretamente da elaboração de textos, imagens, sons, vídeos, argumentos e demais formas de representação simbólica. Essa capacidade amplia extraordinariamente as possibilidades de produção do conhecimento, da criatividade e da inovação, mas também redefine os desafios relacionados à autoria, à confiabilidade da informação, à responsabilidade ética e à construção social da verdade. É precisamente essa convergência entre capacidade generativa, plataformas digitais e reorganização da esfera informacional que permitirá

compreender, na sequência desta discussão, por que as novas inteligências artificiais assumem papel central na amplificação contemporânea dos fenômenos associados à pós-verdade.

A compreensão dos efeitos sociais das inteligências artificiais generativas exige reconhecer que sua influência não decorre apenas da capacidade de produzir conteúdos sintéticos, mas da convergência entre essa capacidade produtiva e as infraestruturas digitais responsáveis por sua circulação. Conforme demonstrado na seção anterior, as plataformas digitais transformaram-se em ecossistemas algorítmicos orientados pela coleta de dados, personalização da informação e maximização do engajamento. As inteligências artificiais generativas inserem-se nesse ambiente ampliando exponencialmente o volume, a velocidade e a sofisticação dos conteúdos que podem ser produzidos e distribuídos. Pela primeira vez na história da comunicação humana, sistemas automatizados são capazes de produzir, adaptar e personalizar textos, imagens, áudios e vídeos em escala compatível com a própria infraestrutura global de circulação informacional. A consequência não consiste apenas no aumento quantitativo da informação disponível, mas na transformação qualitativa das formas pelas quais o conhecimento é produzido, legitimado e compartilhado (UNESCO, 2024; OECD, 2024).

Essa convergência altera profundamente a ecologia informacional contemporânea. Durante grande parte da história da comunicação, produzir conteúdos em larga escala exigia recursos financeiros, competências técnicas e estruturas organizacionais relativamente complexas. A imprensa, o rádio, a televisão e mesmo os primeiros ambientes digitais dependiam da atuação de profissionais especializados e instituições capazes de sustentar processos contínuos de produção. As inteligências artificiais generativas reduzem significativamente essas barreiras ao permitir que indivíduos ou pequenas organizações produzam rapidamente materiais anteriormente restritos a equipes especializadas. A democratização da capacidade produtiva amplia oportunidades criativas e educacionais, mas também reduz os custos associados à fabricação sistemática de desinformação, manipulação audiovisual e campanhas coordenadas de influência (Chesney; Citron, 2019; UNESCO, 2024).

Essa transformação é particularmente evidente na produção textual. Modelos contemporâneos são capazes de elaborar notícias, artigos, comentá-

rios, mensagens em redes sociais, respostas automatizadas, campanhas publicitárias e documentos administrativos em poucos segundos, adaptando linguagem, extensão e estilo ao contexto solicitado. O mesmo princípio aplica-se à geração de imagens, vídeos e vozes sintéticas. Uma campanha de desinformação que anteriormente exigiria grande quantidade de redatores, designers e produtores audiovisuais pode hoje ser parcialmente automatizada mediante integração entre diferentes sistemas generativos. O desafio contemporâneo desloca-se, portanto, da simples produção de conteúdos para a capacidade de distinguir, contextualizar e validar sua origem, seus objetivos e sua confiabilidade (Bender et al., 2021; Chesney; Citron, 2019).

Entretanto, seria metodologicamente inadequado concluir que as inteligências artificiais generativas produzem necessariamente desinformação. A literatura especializada demonstra precisamente o contrário: essas tecnologias apresentam elevado potencial para fortalecer a produção científica, ampliar a acessibilidade, favorecer processos educativos, acelerar pesquisas, democratizar recursos comunicacionais e apoiar atividades profissionais altamente qualificadas. O problema central não reside na existência da tecnologia em si, mas nas condições sociais, econômicas, políticas e comunicacionais que orientam sua utilização. Tal como ocorreu historicamente com outras tecnologias de informação, os efeitos sociais da IA generativa dependem da interação entre arquitetura técnica, modelos de governança, interesses institucionais, regulações jurídicas, competências dos usuários e valores culturais presentes em cada sociedade (UNESCO, 2021; Floridi; Chiriatti, 2020).

Essa constatação conduz a uma distinção conceitual importante entre potencial tecnológico e efeito social. O potencial refere-se às capacidades objetivamente disponibilizadas pelos sistemas generativos; os efeitos dizem respeito às formas concretas pelas quais essas capacidades são apropriadas pelos diferentes atores sociais. Uma mesma ferramenta pode ser utilizada para produzir material didático acessível a estudantes com deficiência, apoiar pesquisadores na organização de grandes bases documentais, auxiliar profissionais da saúde na elaboração de registros clínicos ou fabricar campanhas coordenadas de manipulação política. A análise científica das inteligências artificiais generativas exige, portanto, abandonar interpretações essencialistas

que as classificam como intrinsecamente benéficas ou prejudiciais. Seus impactos emergem da articulação entre tecnologia e contexto social (Russell; Norvig, 2021; UNESCO, 2021).

No âmbito da comunicação pública, contudo, a capacidade de produzir conteúdos sintéticos em grande escala introduz desafios inéditos para os regimes contemporâneos de verdade. A tradição jornalística, científica e documental sempre enfrentou problemas relacionados à falsificação, propaganda e manipulação informacional. O que se modifica no contexto atual é a velocidade, a personalização e o custo de produção dessas estratégias. Sistemas generativos permitem adaptar mensagens para públicos específicos, produzir múltiplas versões de um mesmo conteúdo, alterar estilos discursivos, simular identidades e combinar texto, imagem, voz e vídeo em narrativas altamente convincentes. A produção sintética deixa de ser exceção para tornar-se possibilidade estrutural da comunicação digital (Vosoughi; Roy; Aral, 2018; Chesney; Citron, 2019).

Essa transformação torna particularmente complexa a relação entre autenticidade e credibilidade. Durante longo período, a verificação da autoria e da origem de documentos constituiu elemento central para avaliação de sua confiabilidade. No contexto das inteligências artificiais generativas, entretanto, a autenticidade formal de um conteúdo deixa progressivamente de constituir garantia suficiente de sua veracidade. Um texto perfeitamente estruturado pode conter informações inexistentes; uma fotografia aparentemente documental pode jamais ter correspondido a qualquer acontecimento real; uma gravação de voz pode reproduzir fielmente características de uma pessoa sem que ela jamais tenha pronunciado aquelas palavras. A questão central desloca-se da análise exclusiva do documento para processos mais amplos de validação contextual, rastreabilidade das fontes e verificação independente das informações (UNESCO, 2024; Pasquale, 2015).

Nesse cenário, ganha relevância crescente o conceito de alfabetização em inteligência artificial (*AI literacy*). Assim como a alfabetização midiática tornou-se fundamental para a compreensão crítica dos meios de comunicação, pesquisadores vêm defendendo que sociedades contemporâneas necessitam desenvolver competências específicas para compreender limites, possibilidades e riscos associados aos sistemas generativos. Tais competências incluem reconhecer que

modelos produzem respostas probabilísticas, compreender a importância da verificação de fontes, identificar possíveis vieses, interpretar criticamente conteúdos sintéticos e utilizar essas ferramentas de maneira ética e responsável. A alfabetização em IA não corresponde apenas ao domínio técnico da tecnologia, mas à formação de cidadãos capazes de compreender suas implicações epistemológicas, sociais e políticas (Long; Magerko, 2020; UNESCO, 2024).

Outro aspecto relevante refere-se à transformação da própria produção do conhecimento. Tradicionalmente, a construção científica fundamenta-se na explicitação dos procedimentos metodológicos, na rastreabilidade das fontes, na possibilidade de reprodução dos resultados e na revisão crítica pelos pares. Os modelos generativos introduzem novas mediações nesse processo, uma vez que passam a participar da elaboração de hipóteses, sínteses bibliográficas, estruturas argumentativas e versões preliminares de textos científicos. Essa participação não elimina os princípios epistemológicos da ciência, mas exige novas práticas de transparência quanto ao uso dessas tecnologias, reforçando a necessidade de explicitação metodológica, validação empírica e responsabilidade intelectual dos pesquisadores sobre todas as informações apresentadas (COPE, 2023; Nature, 2023).

Essas discussões revelam que as inteligências artificiais generativas não podem ser compreendidas apenas como instrumentos tecnológicos isolados. Elas integram um processo histórico mais amplo de reorganização das formas pelas quais a sociedade produz, distribui, valida e consome informações. A passagem da comunicação predominantemente humana para ecossistemas híbridos, nos quais agentes humanos e sistemas computacionais participam conjuntamente da elaboração de conteúdos, inaugura novos desafios para a Filosofia, a Sociologia, a Ciência da Informação, a Educação, o Direito e as Ciências Cognitivas. A questão deixa de ser apenas “o que a máquina consegue produzir?” para tornar-se “como a presença dessas máquinas transforma os próprios critérios sociais de produção da verdade?” (Floridi, 2014; Floridi; Chiriatti, 2020).

Sob essa perspectiva, a emergência das inteligências artificiais generativas representa continuidade e, simultaneamente, aprofundamento das transformações analisadas nas seções anteriores desta pesquisa. A pós-verdade evidenciou o enfraquecimento relativo da autoridade exclusiva dos fatos diante da crescen-

te influência de emoções, identidades e crenças. As plataformas digitais reorganizaram a circulação da informação por meio da personalização algorítmica, da economia da atenção e da dataficação. As inteligências artificiais generativas acrescentam um terceiro elemento a esse processo: a automatização da própria produção dos conteúdos que circulam nesses ambientes. Produção, distribuição, recomendação e personalização passam a integrar um mesmo ecossistema computacional, modificando profundamente a arquitetura comunicacional da sociedade contemporânea.

Dessa forma, compreender os diferentes efeitos da pós-verdade gerados e impulsionados pelas redes sociais e pelas novas inteligências artificiais generativas exige abandonar análises fragmentadas dessas tecnologias. Não se trata de fenômenos independentes, mas de dimensões complementares de uma mesma transformação histórica. As plataformas digitais fornecem a infraestrutura global de circulação; os algoritmos organizam visibilidade e engajamento; os modelos generativos ampliam extraordinariamente a capacidade de produção simbólica; e os processos sociais de polarização, disputa identitária e crise de confiança institucional oferecem o contexto no qual esses recursos adquirem significado político e cultural. É precisamente a convergência entre esses elementos que caracteriza a nova ecologia informacional do século XXI e fundamenta a hipótese central desta pesquisa: a de que a pós-verdade contemporânea não pode mais ser compreendida apenas como fenômeno discursivo ou comunicacional, mas como resultado da interação entre infraestruturas sociotécnicas, inteligência artificial generativa e transformações profundas nos regimes de produção e legitimação do conhecimento.

A trajetória histórica reconstruída ao longo desta seção demonstra que as inteligências artificiais generativas não constituem uma ruptura isolada na história da computação, mas o resultado de um longo processo de amadurecimento científico iniciado nas primeiras investigações sobre inteligência artificial, aprendizagem de máquina e modelagem estatística da informação. Desde as abordagens simbólicas desenvolvidas na segunda metade do século XX, passando pela consolidação das redes neurais artificiais, pela expansão do *deep learning*, pelo desenvolvimento das redes generativas adversariais e da arquitetura *Transformer*, até a emergência dos grandes modelos fundacionais, observa-se

um movimento contínuo de ampliação da capacidade computacional de reconhecer padrões, modelar distribuições probabilísticas e produzir representações sintéticas cada vez mais sofisticadas. A aparente novidade das atuais inteligências artificiais generativas decorre, sobretudo, da convergência entre avanços científicos acumulados durante décadas e a disponibilidade contemporânea de grandes volumes de dados, infraestrutura computacional de alta performance e arquiteturas neurais capazes de operar em escala global (Goodfellow; Bengio; Courville, 2016; Russell; Norvig, 2021).

Essa reconstrução histórica permite compreender que as inteligências artificiais generativas não devem ser definidas exclusivamente por sua capacidade de produzir textos, imagens ou outros artefatos digitais, mas por sua inserção em uma nova ecologia da produção do conhecimento. Diferentemente das tecnologias digitais que caracterizaram a primeira fase da internet, cuja principal função consistia em ampliar o acesso à informação previamente existente, os sistemas generativos participam diretamente da elaboração de novos conteúdos, reorganizando as relações entre autoria, criatividade, linguagem, representação e produção intelectual. A máquina deixa de desempenhar apenas funções instrumentais de armazenamento, busca ou processamento de dados para ocupar posição ativa na construção das representações simbólicas que passam a circular socialmente. Essa mudança altera profundamente a própria natureza da mediação tecnológica, deslocando-a da gestão da informação para sua produção contínua (Bommasani et al., 2021; UNESCO, 2024).

Entretanto, conforme demonstrado ao longo deste ensaio, a produção sintética realizada por esses modelos não pode ser confundida com compreensão, consciência ou intencionalidade nos moldes da experiência humana. Os sistemas generativos operam mediante inferências probabilísticas construídas a partir das regularidades presentes nos dados utilizados em seu treinamento. Sua impressionante competência linguística e multimodal decorre da capacidade de modelar distribuições estatísticas extremamente complexas, e não da existência de compreensão semântica, julgamento moral ou experiência subjetiva do mundo. Essa distinção possui importância epistemológica fundamental, pois impede tanto interpretações tecnofóbicas, que atribuem às máquinas capacidades inexistentes, quanto perspectivas excessivamente otimistas que projetam

sobre elas formas de inteligência equivalentes à cognição humana. O reconhecimento dessa especificidade constitui condição indispensável para utilização crítica, ética e cientificamente responsável dessas tecnologias (Bender et al., 2021; Floridi; Chiriatti, 2020).

Ao mesmo tempo, a ausência de consciência não reduz a magnitude dos efeitos sociais produzidos pelas inteligências artificiais generativas. Pelo contrário, sua influência decorre precisamente da capacidade de produzir conteúdos linguisticamente coerentes, visualmente convincentes e comunicacionalmente eficazes em escala sem precedentes. A incorporação desses sistemas aos ambientes educacionais, científicos, profissionais, administrativos e culturais inaugura novas possibilidades de ampliação da produtividade intelectual, democratização do acesso ao conhecimento, personalização da aprendizagem, aceleração da pesquisa científica e fortalecimento da inovação tecnológica. Simultaneamente, impõe desafios inéditos relacionados à autoria, à propriedade intelectual, aos vieses algorítmicos, à concentração tecnológica, à transparência dos modelos e à preservação da autonomia cognitiva dos sujeitos. As inteligências artificiais generativas revelam-se, assim, tecnologias profundamente ambivalentes, cujos efeitos dependem das formas sociais, políticas e éticas pelas quais são incorporadas às práticas humanas (UNESCO, 2021; OECD, 2024).

Essa ambivalência torna-se ainda mais evidente quando os modelos generativos são analisados em articulação com as plataformas digitais discutidas na seção anterior. Isoladamente, a capacidade de produzir conteúdos sintéticos representa uma inovação tecnológica de grande alcance. Integrada às infraestruturas algorítmicas responsáveis pela personalização da informação, pela economia da atenção e pela circulação global de conteúdos, essa mesma capacidade adquire dimensão estrutural. Produção, recomendação, distribuição e adaptação de informações passam a ocorrer dentro de um único ecossistema sociotécnico, no qual algoritmos, plataformas e modelos generativos operam de forma complementar. A inteligência artificial deixa de constituir apenas ferramenta especializada para transformar-se em componente permanente da arquitetura comunicacional contemporânea, influenciando diretamente os processos pelos quais indivíduos constroem conhecimentos, estabelecem vínculos sociais e interpretam os acontecimentos do mundo (Van Dijck; Poell; De Waal, 2018; Zuboff, 2019).

Sob essa perspectiva, a contribuição das inteligências artificiais generativas para a problemática da pós-verdade ultrapassa amplamente a possibilidade de produzir conteúdos falsos ou manipulados. Seu impacto mais profundo reside na transformação das condições estruturais de produção da informação. A redução drástica dos custos de geração de textos, imagens, vídeos e áudios, associada à capacidade de personalização em larga escala, modifica significativamente a dinâmica dos processos comunicacionais contemporâneos. A fronteira entre conteúdos produzidos por seres humanos e por sistemas computacionais torna-se progressivamente menos evidente; a velocidade de elaboração supera amplamente a capacidade tradicional de verificação; e a produção simbólica passa a integrar cadeias automatizadas de circulação orientadas por algoritmos de recomendação e modelos econômicos baseados na captura da atenção. A pós-verdade, nesse contexto, deixa de depender exclusivamente da circulação de informações falsas e passa a relacionar-se também com a crescente dificuldade de estabelecer critérios compartilhados para identificar autoria, confiabilidade, autenticidade e legitimidade dos conteúdos que compõem o espaço público digital (Mcintyre, 2018; UNESCO, 2024).

Dessa forma, as inteligências artificiais generativas não representam apenas uma nova etapa da evolução tecnológica, mas uma inflexão epistemológica na história da produção do conhecimento. Ao automatizar parte significativa da elaboração de representações simbólicas, esses sistemas passam a participar diretamente dos processos sociais de construção da realidade, modificando as relações entre informação, linguagem, verdade e poder. Compreender seus fundamentos técnicos, seus limites cognitivos, suas potencialidades sociais e seus efeitos sobre a circulação contemporânea do conhecimento torna-se, portanto, condição indispensável para analisar criticamente os diferentes efeitos da pós-verdade impulsionados pelas redes sociais e pelas novas inteligências artificiais generativas. É precisamente nessa convergência entre plataformas digitais, sistemas generativos e transformações dos regimes de verdade que se situa o problema central desta pesquisa, cuja discussão será aprofundada nas seções subsequentes por meio da análise integrada desses fenômenos no contexto da sociedade informacional do século XXI.

REFERÊNCIAS

BENDER, Emily M.; GEBRU, Timnit; McMILLAN-MAJOR, Angelina; SHMITCHELL, Shmargaret. On the dangers of stochastic parrots: Can language models be too big? In: **ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY (FAccT '21)**, 2021. **Proceedings [...]**. New York: ACM, 2021. p. 610–623. DOI: <https://doi.org/10.1145/3442188.3445922>. Disponível em: <https://doi.org/10.1145/3442188.3445922>. Acesso em: 14 jul. 2026.

BODEN, Margaret A. **The creative mind: myths and mechanisms**. 2. ed. London: Routledge, 2004.

BOMMASANI, Rishi et al. **On the opportunities and risks of foundation models**. Stanford: Center for Research on Foundation Models (CRFM), Stanford University, 2021. Disponível em: <https://arxiv.org/abs/2108.07258>. Acesso em: 14 jul. 2026.

BROWN, Tom B. et al. Language models are few-shot learners. **Advances in Neural Information Processing Systems**, Red Hook, v. 33, p. 1877–1901, 2020. Disponível em: <https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bf-cb4967418bfb8ac142f64a-Abstract.html>. Acesso em: 14 jul. 2026.

BRYNJOLFSSON, Erik; McAfee, Andrew. **Machine, platform, crowd: harnessing our digital future**. New York: W. W. Norton, 2017.

CHESNEY, Robert; CITRON, Danielle Keats. Deep fakes: a looming challenge for privacy, democracy, and national security. **California Law Review**, Berkeley, v. 107, n. 6, p. 1753–1820, 2019. DOI: <https://doi.org/10.15779/Z38RV0D15J>. Disponível em: <https://doi.org/10.15779/Z38RV0D15J>. Acesso em: 14 jul. 2026.

COMMITTEE ON PUBLICATION ETHICS (COPE). **COPE Position Statement: Authorship and AI Tools**. London: COPE, 2023. Disponível em: <https://publicationethics.org/resources/cope-position-statements/ai-author>. Acesso em: 14 jul. 2026.

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

FLORIDI, Luciano; CHIRIATTI, Massimo. GPT-3: Its nature, scope, limits, and consequences. **Minds and Machines**, Dordrecht, v. 30, n. 4, p. 681–694, 2020. DOI: <https://doi.org/10.1007/s11023-020-09548-1>. Disponível em: <https://doi.org/10.1007/s11023-020-09548-1>. Acesso em: 14 jul. 2026.

GOODFELLOW, Ian et al. Generative adversarial nets. In: **ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS (NIPS 2014)**. **Proceedings [...]**. Montréal, 2014. p. 2672–2680. Disponível em: https://papers.nips.cc/paper_files/paper/2014/hash/f033ed80deb0234979a61f95710dbe25-Abstract.html. Acesso em: 14 jul. 2026.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep learning**. Cambridge: MIT Press, 2016.

HOCHREITER, Sepp; SCHMIDHUBER, Jürgen. Long short-term memory. **Neural Computation**, Cambridge, v. 9, n. 8, p. 1735–1780, 1997. DOI: <https://doi.org/10.1162/neco.1997.9.8.1735>. Disponível em: <https://doi.org/10.1162/neco.1997.9.8.1735>. Acesso em: 14 jul. 2026.

HOLMES, Wayne; BIALIK, Maya; FADEL, Charles. **Artificial intelligence in education: promises and implications for teaching and learning**. Boston: Center for Curriculum Redesign, 2019. Disponível em: <https://curriculumredesign.org/our-work/artificial-intelligence-in-education/>. Acesso em: 14 jul. 2026.

JUMPER, John et al. Highly accurate protein structure prediction with AlphaFold. **Nature**, London, v. 596, n. 7873, p. 583–589, 2021. DOI: <https://doi.org/10.1038/s41586-021-03819-2>. Disponível em: <https://doi.org/10.1038/s41586-021-03819-2>. Acesso em: 14 jul. 2026.

LECUN, Yann; BENGIO, Yoshua; HINTON, Geoffrey. Deep learning. **Nature**, London, v. 521, n. 7553, p. 436–444, 2015. DOI: <https://doi.org/10.1038/nature14539>. Disponível em: <https://doi.org/10.1038/nature14539>. Acesso em: 14 jul. 2026.

LONG, Duri; MAGERKO, Brian. What is AI literacy? Competencies and design considerations. In: **CHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS (CHI '20)**, 2020. **Proceedings [...]**. New York: ACM, 2020. Article 1, 16 p. DOI: <https://doi.org/10.1145/3313831.3376727>. Disponível em: <https://doi.org/10.1145/3313831.3376727>. Acesso em: 14 jul. 2026.

MIAO, Fengchun; HOLMES, Wayne. **Guidance for generative AI in education and research**. Paris: UNESCO, 2023. ISBN 978-92-3-100612-8. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000386693>. Acesso em: 15 jul. 2026.

NATURE. Tools such as ChatGPT threaten transparent science; here are our ground rules for their use. **Nature**, London, v. 613, n. 7945, p. 612, 2023. DOI: <https://doi.org/10.1038/d41586-023-00191-1>. Disponível em: <https://doi.org/10.1038/d41586-023-00191-1>. Acesso em: 14 jul. 2026.

OECD. **Generative AI**. Paris: Organisation for Economic Co-operation and Development, 2024. Disponível em: <https://www.oecd.org/en/topics/generative-ai.html>. Acesso em: 14 jul. 2026.

OPENAI. **GPT-4 Technical Report**. San Francisco, CA: OpenAI, 2023. DOI: <https://doi.org/10.48550/arXiv.2303.08774>. Disponível em: <https://arxiv.org/abs/2303.08774>. Acesso em: 14 jul. 2026.

OUYANG, Long et al. Training language models to follow instructions with human feedback. In: **ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS (NeurIPS 2022). Proceedings [...]**. Red Hook, NY: Curran Associates, 2022. p. 27730–27744. Disponível em: https://proceedings.neurips.cc/paper_files/paper/2022/hash/b1efde53be364a73914f58805a001731-Abstract.html. Acesso em: 14 jul. 2026.

ROMBACH, Robin et al. High-resolution image synthesis with latent diffusion models. In: **IEEE/CVF CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR 2022)**, 2022, New Orleans. **Proceedings [...]**. New York: IEEE, 2022. p. 10684–10695. DOI: <https://doi.org/10.1109/CVPR52688.2022.01042>. Disponível em: <https://doi.org/10.1109/CVPR52688.2022.01042>. Acesso em: 14 jul. 2026.

RUSSELL, Stuart; NORVIG, Peter. **Artificial intelligence: a modern approach**. 4. ed. Hoboken: Pearson, 2021.

UNESCO. **Recommendation on the Ethics of Artificial Intelligence**. Paris: UNESCO, 2021. Disponível em: <https://unesdoc.UNESCO.org/ark:/48223/pf0000381137>. Acesso em: 14 jul. 2026.

UNESCO. **Guidance for Generative AI in Education and Research**. Paris: UNESCO, 2023. Disponível em: <https://unesdoc.UNESCO.org/ark:/48223/pf0000386693>. Acesso em: 14 jul. 2026.

VASWANI, Ashish et al. Attention is all you need. In: **ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS (NIPS 2017)**. Red Hook, NY: Curran Associates, 2017. v. 30, p. 5998–6008. Disponível em: https://papers.nips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html. Acesso em: 14 jul. 2026.

WEIZENBAUM, Joseph. ELIZA: A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, New York, v. 9, n. 1, p. 36–45, 1966.

WESTERLUND, Mika. The emergence of deepfake technology: a review. **Technology Innovation Management Review**, Ottawa, v. 9, n. 11, p. 39–52, 2019. DOI: <https://doi.org/10.22215/timreview/1282>. Disponível em: <https://doi.org/10.22215/timreview/1282>. Acesso em: 14 jul. 2026.

WORLD INTELLECTUAL PROPERTY ORGANIZATION (WIPO). **Generative AI: Navigating Intellectual Property**. Geneva: WIPO, 2024. Disponível em: <https://www.wipo.int/publications/en/details.jsp?id=4713>. Acesso em: 14 jul. 2026.

CAPÍTULO 4

A CONVERGÊNCIA ENTRE PÓS-VERDADE, PLATAFORMAS DIGITAIS E INTELIGÊNCIAS ARTIFICIAIS GENERATIVAS

A reconstrução histórica desenvolvida nos capítulos anteriores evidencia que pós-verdade, plataformas digitais e inteligências artificiais generativas constituem fenômenos conceitualmente distintos, surgidos em momentos históricos diferentes e originalmente investigados por campos científicos relativamente independentes. A pós-verdade emerge como categoria analítica da Filosofia Política, da Epistemologia e dos estudos sobre comunicação, procurando compreender as transformações contemporâneas dos regimes de produção da verdade. As redes sociais desenvolvem-se inicialmente como objeto da Sociologia, consolidando-se posteriormente como plataformas digitais estudadas pelas Ciências da Comunicação, pela Ciência da Informação e pelos Estudos de Plataforma. As inteligências artificiais generativas, por sua vez, possuem origem na Ciência da Computação, na Inteligência Artificial e na Aprendizagem de Máquina. Durante muitos anos, essas três trajetórias evoluíram de maneira relativamente paralela. Entretanto, a segunda década do século XXI inaugura um cenário no qual tais processos deixam de atuar isoladamente e passam a constituir dimensões interdependentes de uma mesma transformação sociotécnica. Compreender essa convergência representa condição indispensável para interpretar a reorganização contemporânea da informação, da comunicação e da própria produção social da verdade (Castells, 2009; McIntyre, 2018; Van Dijck; Poell; De Waal, 2018).

A principal hipótese desenvolvida neste capítulo sustenta que a pós-verdade contemporânea não pode mais ser compreendida exclusivamente como consequência da circulação de informações falsas, da polarização política ou da expansão das redes sociais. Tais elementos permanecem relevantes, mas tornam-se insuficientes diante da profunda reorganização da infraestrutura comunicacional produzida pela articulação entre plataformas digitais, sistemas algorítmicos e inteligências artificiais generativas. A pós-verdade deixa progressivamente de representar apenas uma característica dos discursos públicos para constituir uma

propriedade emergente de ecossistemas sociotécnicos capazes de produzir, selecionar, personalizar, distribuir e amplificar conteúdos em escala planetária. O fenômeno desloca-se, portanto, do plano exclusivamente discursivo para uma dimensão estrutural, na qual arquitetura tecnológica, modelos econômicos, algoritmos, dados e práticas sociais participam conjuntamente da construção das condições contemporâneas de produção do conhecimento (Zuboff, 2019; Srnicek, 2017; Floridi, 2014).

Essa mudança exige reconsiderar a própria noção de verdade utilizada ao longo da história da comunicação. Durante séculos, os processos de validação do conhecimento estiveram fortemente associados à atuação de instituições sociais relativamente estáveis, como universidades, periódicos científicos, sistemas educacionais, editoras, organizações jornalísticas e órgãos públicos responsáveis pela produção e certificação de informações. Embora tais instituições jamais tenham eliminado disputas políticas, ideológicas ou econômicas, elas constituíam espaços relativamente reconhecidos para a verificação dos fatos, o confronto entre interpretações concorrentes e a produção de consensos provisórios acerca da realidade. A autoridade epistemológica encontrava-se distribuída entre organizações especializadas cuja legitimidade derivava, em grande medida, da existência de procedimentos institucionalizados de validação do conhecimento (Arendt, 1967; Foucault, 1979; McIntyre, 2018).

O advento da internet inicia um processo de descentralização desses mecanismos de mediação. Conforme discutido anteriormente, a Web amplia extraordinariamente a capacidade de indivíduos produzirem e compartilharem informações sem depender da autorização prévia de instituições tradicionais. Essa transformação representa importante conquista democrática, pois diversifica vozes historicamente excluídas da esfera pública e amplia significativamente o acesso à produção de conhecimento. Entretanto, a mesma infraestrutura que democratiza a emissão de conteúdos reduz a centralidade dos antigos mecanismos institucionais de filtragem e validação. A emergência das plataformas digitais aprofunda ainda mais esse movimento ao substituir parte da mediação editorial por sistemas algorítmicos cuja lógica de funcionamento se orienta predominantemente por critérios relacionados ao engajamento, à personalização e à retenção da atenção dos usuários (Castells, 2009; Gillespie, 2018; Van Dijck, 2013).

Essa reorganização da mediação informacional constitui o primeiro eixo da convergência aqui proposta. As plataformas digitais deixam de desempenhar apenas a função de espaços de encontro entre indivíduos para converterem-se em infraestruturas responsáveis por organizar a própria experiência informacional da sociedade. A informação não circula mais simplesmente porque foi produzida; ela circula porque algoritmos decidem, continuamente, quais conteúdos serão priorizados, para quem serão apresentados, em que momento aparecerão e com quais outros conteúdos estabelecerão relações de proximidade. A arquitetura algorítmica passa a integrar o próprio processo de construção da realidade social, uma vez que influencia significativamente aquilo que indivíduos, grupos e instituições conseguem perceber como relevante em determinado contexto histórico (Gillespie, 2018; Van Dijck; Poell; De Waal, 2018).

Entretanto, a reorganização da circulação da informação representa apenas parte da transformação contemporânea. O segundo eixo dessa convergência emerge quando as plataformas passam a operar simultaneamente com sistemas de inteligência artificial capazes não apenas de recomendar conteúdos existentes, mas também de produzir novos conteúdos em tempo real. A inteligência artificial generativa altera profundamente o equilíbrio histórico entre produção e distribuição da informação. Durante décadas, as plataformas concentraram-se prioritariamente na organização de conteúdos produzidos por seres humanos. A incorporação dos grandes modelos generativos modifica essa lógica ao permitir que o próprio ecossistema digital participe diretamente da elaboração de textos, imagens, vídeos, vozes e demais representações simbólicas que posteriormente serão distribuídas pelas plataformas. Produção e circulação deixam de constituir processos relativamente independentes para integrar uma mesma infraestrutura computacional (OpenAI, 2023; UNESCO, 2024).

Essa convergência inaugura aquilo que se pode denominar produção algorítmica da realidade simbólica. Evidentemente, as inteligências artificiais não produzem a realidade em sentido ontológico; fatos continuam ocorrendo independentemente dos modelos computacionais. O que se transforma é a produção das representações mediante as quais a sociedade interpreta, organiza e atribui significado aos acontecimentos. A realidade social sempre dependeu de processos simbólicos de descrição, interpretação e narrativa. A novidade contemporânea

nea consiste no fato de que parte crescente dessas narrativas passa a ser elaborada por sistemas estatísticos capazes de operar em velocidade, escala e grau de personalização impossíveis para a comunicação exclusivamente humana. A consequência não reside apenas na multiplicação dos conteúdos disponíveis, mas na modificação das próprias condições estruturais da produção simbólica (Floridi, 2014; Floridi; Chiriatti, 2020).

Sob essa perspectiva, a inteligência artificial generativa não substitui as plataformas digitais; ela amplia significativamente suas capacidades estruturantes. Enquanto os algoritmos de recomendação selecionam quais conteúdos alcançarão maior visibilidade, os modelos generativos tornam possível produzir rapidamente versões específicas desses conteúdos para diferentes públicos, idiomas, contextos culturais e perfis comportamentais. A personalização deixa de atuar apenas sobre a distribuição da informação para alcançar igualmente sua produção. Cada indivíduo pode, em princípio, receber conteúdos adaptados às suas preferências cognitivas, emocionais, políticas ou culturais, ampliando consideravelmente a capacidade das plataformas de modular experiências informacionais individualizadas. A informação torna-se simultaneamente personalizada em sua circulação e em sua própria elaboração, característica inédita na história da comunicação humana (Zuboff, 2019; Bommasani et al., 2021).

É precisamente nesse ponto que a problemática da pós-verdade assume nova complexidade. Se, anteriormente, o desafio consistia em compreender por que determinadas informações falsas circulavam mais rapidamente do que informações verificadas, passa a ser necessário investigar como sistemas computacionais participam da própria arquitetura de produção dessas narrativas. A distinção tradicional entre emissor, meio e receptor torna-se progressivamente insuficiente para explicar ambientes nos quais algoritmos recomendam conteúdos, modelos generativos os produzem, plataformas os distribuem e usuários simultaneamente os consomem, reinterpretam, modificam e redistribuem. A comunicação contemporânea deixa de ser predominantemente linear para constituir um sistema adaptativo complexo, caracterizado pela interação contínua entre agentes humanos e agentes computacionais (Castells, 2009; Floridi, 2014).

Essa mudança representa uma inflexão epistemológica de grande alcance. Durante séculos, as principais discussões sobre verdade concentraram-se na cor-

respondência entre afirmações e fatos, na coerência lógica dos argumentos ou na legitimidade dos procedimentos científicos de validação do conhecimento. Embora essas questões permaneçam fundamentais, a sociedade informacional acrescenta uma nova dimensão ao problema: as condições infraestruturais pelas quais determinadas representações tornam-se socialmente visíveis, compartilháveis e influentes. A verdade deixa de depender exclusivamente da qualidade das evidências apresentadas e passa a ser profundamente condicionada pelas arquiteturas tecnológicas responsáveis pela organização dos fluxos informacionais. Nesse sentido, compreender a pós-verdade contemporânea exige investigar não apenas os conteúdos das mensagens, mas também os sistemas que regulam sua produção, circulação e permanência na esfera pública digital.

A formulação dessa hipótese permite propor um deslocamento analítico em relação à maior parte da literatura contemporânea. Grande número de estudos dedica-se a investigar separadamente os efeitos das redes sociais sobre a polarização política, os impactos da inteligência artificial sobre o mercado de trabalho ou as implicações epistemológicas da pós-verdade. Embora essas abordagens tenham produzido contribuições fundamentais, elas frequentemente preservam fronteiras disciplinares que dificultam compreender a dinâmica integrada do ecossistema informacional contemporâneo. A presente investigação parte de pressuposto distinto: plataformas digitais, algoritmos de recomendação, inteligência artificial generativa, economia da atenção e pós-verdade não constituem fenômenos paralelos, mas componentes de uma mesma arquitetura sociotécnica. Cada um desses elementos modifica o funcionamento dos demais, produzindo efeitos emergentes que não podem ser explicados isoladamente. O desafio científico desloca-se, portanto, da análise de tecnologias específicas para a compreensão das relações sistêmicas que articulam produção, circulação, legitimação e consumo da informação (Castells, 2009; Van Dijck; Poell; De Waal, 2018; Floridi, 2014).

Sob essa perspectiva, a pós-verdade deixa de representar apenas um problema relacionado ao conteúdo das mensagens para tornar-se um fenômeno associado à reorganização dos regimes contemporâneos de produção da verdade. Inspirando-se na concepção foucaultiana de regimes de verdade, pode-se compreender que cada formação histórica desenvolve mecanismos específicos pelos quais determinadas formas de conhecimento adquirem legitimidade, enquanto

outras permanecem marginalizadas. A verdade nunca circula independentemente das instituições, das tecnologias e das relações de poder que tornam possível sua produção e reconhecimento social. As plataformas digitais e as inteligências artificiais generativas passam a integrar esse conjunto de dispositivos, não substituindo a ciência, a educação ou o jornalismo, mas interferindo significativamente nas condições pelas quais seus discursos alcançam visibilidade, reconhecimento e influência pública. A questão central deixa de ser apenas “o que é verdadeiro?” para incorporar outra pergunta igualmente relevante: “quais infraestruturas tornam determinadas versões da realidade mais visíveis, compartilháveis e socialmente persuasivas?” (Foucault, 1979; Gillespie, 2018).

Essa reorganização pode ser compreendida a partir de três movimentos históricos sucessivos. O primeiro corresponde à democratização da produção de conteúdos promovida pela internet, que ampliou extraordinariamente a capacidade de indivíduos e grupos sociais participarem da esfera pública. O segundo refere-se à plataformização da comunicação, caracterizada pela crescente centralidade de empresas responsáveis por organizar algoritmicamente os fluxos informacionais mediante sistemas de recomendação, personalização e monetização da atenção. O terceiro movimento, atualmente em curso, consiste na automatização parcial da própria produção simbólica por meio das inteligências artificiais generativas. Esses três processos não substituem uns aos outros; acumulam-se historicamente, produzindo um ambiente comunicacional qualitativamente distinto daquele existente nas primeiras décadas da internet. A comunicação digital contemporânea caracteriza-se precisamente pela integração entre produção distribuída, circulação algorítmica e geração automatizada de conteúdos (Srnicek, 2017; Zuboff, 2019; UNESCO, 2024).

A interação entre esses movimentos permite compreender por que os diferentes efeitos da pós-verdade assumem intensidade inédita no contexto atual. Informações falsas, distorcidas ou descontextualizadas sempre existiram. Propaganda política, rumores, boatos, manipulação documental e estratégias de persuasão acompanham praticamente toda a história das sociedades humanas. O diferencial contemporâneo não reside na existência desses fenômenos, mas na infraestrutura tecnológica que potencializa sua produção, adaptação e circulação. As plataformas digitais ampliam exponencialmente o alcance das

mensagens; os algoritmos identificam públicos potencialmente mais receptivos; os modelos generativos reduzem drasticamente os custos de produção de conteúdos personalizados; e a economia da atenção recompensa precisamente os materiais com maior capacidade de mobilização emocional. A convergência desses elementos modifica estruturalmente a ecologia informacional na qual se desenvolvem os processos de formação da opinião pública (Mcintyre, 2018; Vosoughi; Roy; Aral, 2018).

Nesse contexto, torna-se insuficiente compreender a circulação da informação apenas mediante o modelo clássico emissor–mensagem–receptor. Os ecossistemas digitais contemporâneos operam segundo dinâmica muito mais complexa, caracterizada pela existência de múltiplos agentes humanos e não humanos continuamente interagindo. Usuários produzem conteúdos; algoritmos classificam sua relevância; plataformas reorganizam sua visibilidade; modelos generativos produzem novas versões; sistemas automatizados monitoram padrões de engajamento; anunciantes influenciam estratégias de distribuição; e novos ciclos de interação retroalimentam continuamente o sistema. A comunicação deixa de constituir processo linear para assumir características próprias dos sistemas adaptativos complexos, nos quais pequenas alterações locais podem produzir consequências amplificadas em toda a rede (Castells, 2009; Van Dijck; Poell; De Waal, 2018).

É precisamente essa dinâmica que permite propor um modelo integrador para compreender os diferentes efeitos da pós-verdade. Em vez de situar a origem do fenômeno exclusivamente nos indivíduos, nas tecnologias ou nas instituições, o modelo aqui defendido compreende a pós-verdade como resultado da interação permanente entre cinco dimensões estruturais. A primeira corresponde à dimensão cognitiva, relacionada às limitações humanas de processamento da informação, aos vieses cognitivos, às emoções e aos processos de construção identitária. A segunda refere-se à dimensão tecnológica, constituída pelos algoritmos, pelas plataformas digitais e pelos sistemas de inteligência artificial responsáveis pela organização da experiência informacional. A terceira corresponde à dimensão econômica, fundamentada na monetização dos dados, na economia da atenção e nos modelos de negócios das grandes plataformas digitais. A quarta dimensão é política, envolvendo disputas por poder, legitimidade, influência e controle

da informação. Finalmente, a quinta dimensão corresponde ao plano cultural, no qual crenças, valores, narrativas coletivas e formas de pertencimento social atribuem significado às informações produzidas e compartilhadas. Nenhuma dessas dimensões, isoladamente, explica a pós-verdade; é sua interação contínua que produz o fenômeno observado na sociedade contemporânea (Zuboff, 2019; Cesarino, 2021; Floridi, 2014).

A principal contribuição desse modelo consiste em deslocar a análise da busca por causas únicas para a compreensão das relações entre diferentes níveis de organização social. Explicações centradas exclusivamente na irracionalidade dos indivíduos desconsideram a influência exercida pelas arquiteturas algorítmicas. Interpretações que atribuem responsabilidade apenas às plataformas ignoram processos históricos de polarização, desigualdade e crise institucional. Perspectivas tecnodeterministas tendem a reduzir fenômenos sociais complexos a características específicas das tecnologias. Em sentido oposto, abordagens estritamente culturalistas frequentemente subestimam a importância das infraestruturas computacionais que condicionam a circulação contemporânea das informações. O modelo integrador proposto procura justamente superar essas dicotomias, reconhecendo que a pós-verdade constitui fenômeno emergente produzido pela articulação simultânea entre estruturas técnicas, práticas sociais, interesses econômicos e processos culturais (Floridi, 2014; Gillespie, 2018; Zuboff, 2019).

Essa perspectiva também modifica a compreensão da própria inteligência artificial generativa. Em vez de tratá-la apenas como ferramenta de produção textual ou visual, ela passa a ser compreendida como elemento constitutivo de uma nova infraestrutura epistemológica. Ao produzir respostas, sínteses, argumentos, imagens e diferentes formas de representação simbólica, os modelos generativos participam diretamente da mediação entre sujeitos e conhecimento. Eles não apenas auxiliam indivíduos a acessar informações previamente existentes, mas interferem na forma como essas informações são reorganizadas, interpretadas e apresentadas. A inteligência artificial deixa de ocupar posição periférica nos processos de comunicação para tornar-se agente relevante na constituição das narrativas que organizam a experiência social contemporânea. Essa mediação, entretanto, permanece condicionada por escolhas humanas relativas aos dados de treinamento, aos critérios de alinhamento, aos objetivos institucionais e às políti-

cas de governança das plataformas que disponibilizam esses sistemas (Bender et al., 2021; OpenAI, 2023; UNESCO, 2024).

Compreendida dessa maneira, a convergência entre pós-verdade, plataformas digitais e inteligência artificial generativa inaugura aquilo que se pode denominar regime algorítmico de mediação do conhecimento. Nesse regime, a produção, a circulação e a legitimação das informações deixam de depender exclusivamente da atuação direta de instituições humanas para incorporar infraestruturas computacionais capazes de selecionar, organizar, sintetizar e redistribuir conteúdos em tempo real. A verdade continua sendo construída socialmente, mas passa a fazê-lo em ambientes profundamente mediados por sistemas algorítmicos cuja lógica de funcionamento permanece, em grande medida, invisível aos próprios usuários. O desafio contemporâneo deixa, assim, de consistir apenas em distinguir o verdadeiro do falso; torna-se igualmente necessário compreender as condições sociotécnicas que tornam determinadas versões da realidade mais prováveis, mais visíveis e mais influentes do que outras. É nessa transformação que reside, provavelmente, uma das características mais profundas da sociedade informacional do século XXI e o principal eixo interpretativo desenvolvido nesta obra.

A proposição de um regime algorítmico de mediação do conhecimento permite avançar para uma consequência ainda mais profunda da convergência entre plataformas digitais e inteligências artificiais generativas: a transformação da própria natureza da autoridade epistemológica. Tradicionalmente, a legitimidade do conhecimento científico, jornalístico ou institucional estava associada à identificação relativamente clara de seus produtores, aos procedimentos metodológicos empregados e à possibilidade de escrutínio público das evidências apresentadas. Na sociedade informacional contemporânea, entretanto, parcela crescente das interações cognitivas passa a ocorrer por meio de sistemas capazes de sintetizar, reorganizar e produzir respostas sem que o usuário tenha acesso imediato às fontes, aos critérios de seleção ou às etapas intermediárias do processo de construção da informação. A autoridade deixa de derivar exclusivamente da reputação institucional dos emissores e passa a ser compartilhada com infraestruturas computacionais cuja legitimidade repousa, muitas vezes, na própria eficiência percebida de suas respostas (Pasquale, 2015; Floridi, 2014).

Esse deslocamento modifica significativamente a relação entre confiança e conhecimento. Nas sociedades modernas, confiar significava, em grande medida, reconhecer a competência de determinadas instituições especializadas para produzir conhecimentos sobre campos específicos da realidade. A ciência conquistava autoridade mediante procedimentos de validação empírica, revisão por pares e possibilidade permanente de contestação; o jornalismo fundamentava sua legitimidade na investigação dos fatos e na responsabilidade editorial; o sistema educacional organizava a transmissão intergeracional do conhecimento por meio de currículos, especialistas e práticas pedagógicas institucionalizadas. As tecnologias digitais não eliminam essas formas tradicionais de confiança, mas introduzem novos mediadores cuja autoridade decorre menos da transparência de seus procedimentos e mais da experiência cotidiana de eficiência, rapidez e conveniência oferecida aos usuários. A confiança desloca-se progressivamente da instituição para a interface (Giddens, 1991; Pasquale, 2015).

Essa mudança torna-se particularmente evidente na utilização cotidiana dos grandes modelos de linguagem. Milhões de indivíduos recorrem diariamente às inteligências artificiais generativas para esclarecer dúvidas, produzir textos, interpretar documentos, elaborar códigos computacionais, traduzir conteúdos e organizar informações. Em muitos casos, o usuário estabelece uma relação de confiança operacional baseada na qualidade aparente das respostas recebidas, independentemente de conhecer os mecanismos internos pelos quais tais respostas foram produzidas. O modelo computacional assume progressivamente funções anteriormente desempenhadas por mecanismos tradicionais de consulta, síntese e organização do conhecimento. Embora essa transformação amplie significativamente o acesso à informação, ela também desloca parte do processo de validação para sistemas cuja arquitetura permanece opaca para a maioria dos usuários (OpenAI, 2023; UNESCO, 2024).

Essa opacidade constitui uma das características centrais da nova ecologia informacional. Frank Pasquale denomina de “sociedade da caixa-preta” o contexto em que decisões de elevada relevância social passam a ser mediadas por sistemas computacionais cujos critérios internos permanecem inacessíveis ao escrutínio público. O problema não reside apenas na complexidade matemática dos algoritmos, mas na combinação entre segredo comercial, elevada sofistica-

ção técnica e crescente dependência social dessas infraestruturas. À medida que algoritmos e modelos generativos passam a participar da seleção de informações, da recomendação de conteúdos e da produção de respostas, amplia-se a necessidade de mecanismos democráticos de transparência, auditabilidade e responsabilização capazes de preservar a confiança pública sem comprometer a inovação tecnológica (Pasquale, 2015; UNESCO, 2021).

A opacidade algorítmica também modifica profundamente a temporalidade da circulação do conhecimento. Historicamente, a produção científica desenvolveu-se segundo ritmos relativamente lentos de investigação, revisão, publicação, replicação e consolidação dos resultados. As plataformas digitais introduziram uma lógica de atualização permanente, na qual fluxos informacionais são reorganizados continuamente em tempo real. As inteligências artificiais generativas intensificam ainda mais essa aceleração ao produzir instantaneamente respostas para demandas específicas. O conhecimento deixa progressivamente de ser percebido como resultado de processos longos de investigação para assumir a aparência de disponibilidade imediata. Essa mudança temporal influencia expectativas sociais acerca da produção intelectual, favorecendo a valorização crescente da rapidez em detrimento da reflexão prolongada e da análise metodologicamente cuidadosa (Rosa, 2013; Floridi, 2014).

Essa aceleração não produz apenas transformações quantitativas. Ela altera a própria experiência cognitiva dos sujeitos diante da informação. Quando respostas complexas tornam-se imediatamente acessíveis, reduz-se o tempo dedicado à formulação das perguntas, à busca autônoma de fontes, ao confronto entre interpretações concorrentes e à elaboração gradual dos argumentos. Evidentemente, a inteligência artificial pode também favorecer investigações mais sofisticadas ao reduzir o tempo destinado a tarefas repetitivas. Contudo, essa potencialidade somente se realiza quando os sistemas são utilizados como instrumentos de ampliação da reflexão crítica, e não como substitutos do processo intelectual humano. A diferença entre essas duas formas de utilização constitui uma das questões pedagógicas mais relevantes da atualidade (UNESCO, 2023; Holmes; Bialik; Fadel, 2019).

Nesse contexto, emerge um fenômeno que pode ser denominado delegação cognitiva. Ao longo da história, diferentes tecnologias ampliaram capacidades

humanas específicas: a escrita expandiu a memória; a imprensa multiplicou a circulação do conhecimento; os computadores potencializaram cálculos complexos; os mecanismos de busca reorganizaram o acesso à informação. As inteligências artificiais generativas acrescentam uma nova dimensão a esse processo ao assumir parcialmente atividades relacionadas à produção textual, à síntese argumentativa, à organização conceitual e à elaboração de respostas. A delegação deixa de incidir apenas sobre operações mecânicas ou cálculos matemáticos para alcançar funções tradicionalmente associadas à produção simbólica. Essa característica exige reconsiderar cuidadosamente quais competências permanecerão essencialmente humanas e quais poderão ser progressivamente compartilhadas com sistemas computacionais (Floridi; Chiriatti, 2020; Russell; Norvig, 2021).

A delegação cognitiva, entretanto, não implica necessariamente empobrecimento intelectual. Ao longo da história das tecnologias, diferentes instrumentos inicialmente percebidos como ameaças transformaram-se posteriormente em extensões das capacidades humanas. A escrita foi criticada por supostamente enfraquecer a memória; a imprensa foi acusada de banalizar o conhecimento; as calculadoras despertaram receios acerca da perda das habilidades matemáticas; a internet passou a ser associada à superficialidade informacional. Em todos esses casos, as tecnologias efetivamente reorganizaram determinadas competências, ao mesmo tempo em que possibilitaram o desenvolvimento de outras. A inteligência artificial generativa provavelmente seguirá trajetória semelhante. O desafio científico consiste menos em decidir se essas tecnologias devem ou não ser utilizadas e mais em compreender quais capacidades cognitivas precisam ser preservadas, fortalecidas e ressignificadas em um contexto de crescente cooperação entre seres humanos e sistemas inteligentes (Lévy, 1999; Floridi, 2014).

Essa cooperação torna particularmente relevante a distinção entre automação e ampliação cognitiva. A automação busca substituir determinadas tarefas humanas por processos computacionais capazes de executá-las com menor custo ou maior velocidade. A ampliação cognitiva, por outro lado, compreende a inteligência artificial como instrumento destinado a expandir as capacidades analíticas, criativas e interpretativas dos sujeitos. Essa segunda perspectiva aproxima-se da noção de inteligência aumentada (*augmented intelligence*), segundo a qual o objetivo principal das tecnologias inteligentes não consiste

em reproduzir integralmente a inteligência humana, mas em potencializar processos colaborativos entre pessoas e sistemas computacionais. Sob essa ótica, a qualidade da interação depende menos da sofisticação isolada do algoritmo e mais da capacidade dos usuários de formular perguntas pertinentes, interpretar criticamente as respostas e integrá-las a processos reflexivos mais amplos (Floridi; Chiriatti, 2020; UNESCO, 2021).

É justamente nesse ponto que a convergência entre pós-verdade, plataformas digitais e inteligências artificiais generativas assume sua dimensão mais profunda. O problema central deixa de ser apenas tecnológico ou comunicacional para tornar-se essencialmente epistemológico. A sociedade contemporânea passa a enfrentar não somente o desafio de produzir informações verdadeiras, mas também o de construir competências capazes de avaliar criticamente sistemas que participam ativamente da produção, organização e circulação dessas informações. A crise da verdade deixa de corresponder exclusivamente à disputa entre fatos e opiniões para incorporar uma nova questão: *como preservar condições socialmente compartilhadas de confiança, validação e responsabilidade em ecossistemas informacionais cada vez mais mediados por algoritmos e inteligências artificiais?* Essa pergunta sintetiza o núcleo da problemática desenvolvida nesta obra e prepara o caminho para sua conclusão geral, na qual serão discutidas as implicações dessa nova arquitetura sociotécnica para a democracia, a educação, a ciência e a formação da cidadania no século XXI.

A análise desenvolvida até este ponto permite afirmar que a convergência entre pós-verdade, plataformas digitais e inteligências artificiais generativas não representa apenas uma transformação tecnológica, mas uma reconfiguração profunda da própria ecologia do conhecimento. Durante grande parte da modernidade, os processos de produção da verdade organizavam-se segundo uma lógica predominantemente institucional. A escola, a universidade, os periódicos científicos, os livros, os arquivos públicos e o jornalismo constituíam os principais ambientes responsáveis pela produção, validação e circulação do conhecimento socialmente reconhecido. Evidentemente, tais instituições jamais foram neutras ou isentas de disputas políticas, econômicas e culturais. Entretanto, compartilhavam um aspecto fundamental: a existência de procedimentos relativamente explí-

bitos de validação, revisão e responsabilização sobre aquilo que era produzido. A credibilidade do conhecimento encontrava-se vinculada não apenas ao conteúdo das informações, mas também à legitimidade dos processos institucionais que as sustentavam (Arendt, 1967; Giddens, 1991).

A sociedade digital reorganiza profundamente essa arquitetura institucional. O conhecimento passa progressivamente a circular em ambientes cuja principal característica não consiste na estabilidade das informações, mas na permanente atualização dos fluxos comunicacionais. O documento deixa de ocupar posição central como unidade privilegiada da memória social para dividir espaço com conteúdos efêmeros, continuamente reorganizados por algoritmos que priorizam relevância contextual, engajamento e personalização. A lógica documental cede lugar, em parte, à lógica do fluxo. A informação já não permanece organizada prioritariamente segundo critérios cronológicos, bibliográficos ou arquivísticos, mas conforme sistemas computacionais que reconstroem continuamente sua visibilidade em função das interações realizadas pelos próprios usuários. Essa transformação modifica não apenas a circulação da informação, mas também a forma como a sociedade estabelece relações de permanência, autoridade e memória coletiva (Castells, 2009; Van Dijck, 2013).

As inteligências artificiais generativas intensificam significativamente esse processo ao introduzir uma característica inédita na história da comunicação: a produção contínua de conteúdos potencialmente ilimitados. Durante séculos, a escassez constituiu uma característica inerente à produção intelectual. Escrever um livro, produzir um artigo científico, elaborar um relatório técnico ou criar uma obra artística demandava tempo, competências especializadas e recursos materiais relativamente elevados. A inteligência artificial reduz drasticamente esse custo marginal de produção simbólica. Em poucos segundos, torna-se possível gerar centenas de versões de um mesmo texto, milhares de imagens sobre um único tema ou inúmeras adaptações linguísticas para públicos específicos. A escassez informacional, característica de diferentes períodos históricos, cede progressivamente espaço à abundância extrema de conteúdos produzidos por sistemas automatizados (OpenAI, 2023; UNESCO, 2024).

Essa abundância produz um paradoxo central da sociedade informacional. Ao mesmo tempo em que amplia extraordinariamente o acesso ao conhecimen-

to, dificulta crescentemente a identificação das informações mais confiáveis, relevantes e socialmente significativas. Herbert Simon (1971) já havia observado que uma sociedade rica em informação tende inevitavelmente a produzir escassez de atenção. A inteligência artificial generativa amplia ainda mais essa dinâmica ao acelerar exponencialmente a produção de conteúdos que disputam continuamente o tempo cognitivo disponível dos indivíduos. O problema contemporâneo deixa de consistir apenas na obtenção da informação e passa a concentrar-se na seleção, validação, contextualização e interpretação crítica dos enormes fluxos informacionais disponíveis. Em outras palavras, a principal limitação desloca-se da produção para a capacidade humana de atribuir significado à abundância de conteúdos produzidos (Davenport; Beck, 2001).

Esse deslocamento altera profundamente os processos educacionais. Durante longo período, a escola organizou-se em torno da transmissão de conhecimentos relativamente escassos e de difícil acesso. O professor desempenhava papel central como mediador entre estudantes e acervos limitados de informações. Na sociedade contemporânea, a escassez desloca-se para outro plano. O desafio pedagógico deixa de consistir prioritariamente em fornecer acesso ao conhecimento para concentrar-se na formação de sujeitos capazes de interpretar criticamente informações abundantes, identificar evidências confiáveis, reconhecer manipulações discursivas, compreender os limites dos sistemas algorítmicos e utilizar inteligências artificiais de maneira intelectualmente autônoma. A alfabetização digital amplia-se para uma alfabetização epistemológica, na qual compreender como o conhecimento é produzido torna-se tão importante quanto conhecer seus conteúdos (UNESCO, 2023; Holmes; Bialik; Fadel, 2019).

Sob essa perspectiva, a educação assume papel estratégico na preservação das condições democráticas de produção da verdade. Não se trata apenas de ensinar estudantes a utilizar ferramentas digitais ou inteligências artificiais, mas de desenvolver competências relacionadas ao pensamento crítico, à avaliação de evidências, à compreensão das metodologias científicas, à análise de fontes documentais e à reflexão ética sobre o uso das tecnologias. Em um ambiente caracterizado pela crescente automatização da produção simbólica, torna-se insuficiente dominar procedimentos técnicos de acesso à informação; torna-se neces-

sário compreender os mecanismos sociais, políticos, econômicos e tecnológicos responsáveis por sua organização. A formação cidadã passa, assim, a depender da capacidade de interpretar criticamente ecossistemas sociotécnicos complexos, nos quais conhecimento, algoritmos, plataformas e inteligência artificial encontram-se profundamente articulados (UNESCO, 2021; OECD, 2024).

Essa discussão possui implicações igualmente relevantes para a democracia contemporânea. Regimes democráticos dependem da existência de espaços públicos capazes de sustentar debates informados, pluralidade de perspectivas e condições mínimas de confiança entre os diferentes atores sociais. A convergência entre plataformas digitais e inteligências artificiais generativas amplia simultaneamente oportunidades e desafios para esses processos. Por um lado, reduz barreiras de participação pública, democratiza ferramentas de comunicação e fortalece possibilidades de mobilização coletiva. Por outro, amplia significativamente a capacidade de produção automatizada de campanhas de influência, conteúdos manipulados, simulações audiovisuais e estratégias sofisticadas de segmentação discursiva. A defesa da democracia deixa progressivamente de depender apenas da proteção da liberdade de expressão para incorporar também a necessidade de fortalecer a integridade dos ecossistemas informacionais nos quais essa liberdade se exerce (UNESCO, 2021; Chesney; Citron, 2019).

Essa transformação exige igualmente reconsiderar a relação entre regulação e inovação tecnológica. Durante muito tempo, o debate público foi apresentado como oposição entre desenvolvimento tecnológico e intervenção regulatória. A literatura recente, entretanto, demonstra que essa dicotomia se revela insuficiente diante da crescente centralidade das plataformas e das inteligências artificiais para a organização da vida social. A questão deixa de ser simplesmente regular ou não regular tecnologias digitais. O problema desloca-se para a construção de modelos de governança capazes de promover inovação científica, proteger direitos fundamentais, assegurar transparência algorítmica, preservar a diversidade informacional e responsabilizar adequadamente os diferentes atores envolvidos na produção e circulação automatizada do conhecimento. A governança da inteligência artificial torna-se, portanto, elemento constitutivo da própria governança democrática das sociedades digitais (UNESCO, 2021; OECD, 2024).

Diante desse cenário, torna-se possível sintetizar a principal contribuição teórica desenvolvida ao longo desta obra. A pós-verdade contemporânea não

deve ser compreendida como simples consequência da irracionalidade humana, da manipulação política ou da existência das redes sociais. Tampouco resulta exclusivamente do surgimento das inteligências artificiais generativas. O fenômeno emerge da convergência entre transformações cognitivas, econômicas, tecnológicas, culturais e institucionais que reorganizam profundamente a infraestrutura da comunicação contemporânea. Plataformas digitais estruturam a circulação da informação; algoritmos organizam sua visibilidade; modelos generativos ampliam exponencialmente sua produção; economias baseadas na atenção orientam sua monetização; processos identitários atribuem significado às narrativas produzidas; e sujeitos humanos continuam interpretando, compartilhando, contestando e ressignificando essas informações. A pós-verdade constitui precisamente a manifestação mais visível dessa nova arquitetura sociotécnica do conhecimento.

Essa interpretação permite superar explicações simplificadoras frequentemente presentes no debate público. Não se trata de atribuir às tecnologias digitais responsabilidade exclusiva pela crise contemporânea da verdade, nem de reduzir os problemas sociais à ação de indivíduos desinformados ou emocionalmente manipuláveis. A análise desenvolvida demonstra que os diferentes efeitos da pós-verdade emergem de um sistema complexo de interdependências no qual tecnologias, instituições, interesses econômicos, culturas políticas e práticas comunicacionais influenciam-se reciprocamente. A compreensão desse sistema constitui condição indispensável para formular estratégias efetivas de educação, regulação, inovação e fortalecimento democrático compatíveis com os desafios da sociedade informacional do século XXI.

Por essa razão, a investigação realizada ao longo deste livro conduz a uma conclusão fundamental: a discussão sobre pós-verdade deixou de pertencer exclusivamente ao domínio da Filosofia, da Comunicação ou da Ciência Política. Ela tornou-se questão transversal que envolve simultaneamente Epistemologia, Sociologia, Ciência da Informação, Inteligência Artificial, Educação, Direito, Economia Digital e Estudos de Plataforma. Compreender a crise contemporânea da verdade exige compreender também as infraestruturas computacionais que organizam a circulação do conhecimento, as plataformas que estruturam os fluxos informacionais, os algoritmos que modulam a atenção e os modelos generativos que passam a participar diretamente da produção das representações sociais.

É nesse cruzamento interdisciplinar que se encontra não apenas o principal desafio científico da atualidade, mas também uma das tarefas intelectuais mais relevantes para a preservação das condições democráticas de produção, circulação e validação do conhecimento nas sociedades contemporâneas.

REFERÊNCIAS

ARENDT, Hannah. **Truth and politics**. *The New Yorker*, New York, 25 fev. 1967. Posteriormente republicado em: ARENDT, Hannah. *Between Past and Future: Eight Exercises in Political Thought*. New York: Viking Press, 1968.

BENDER, Emily M.; GEBRU, Timnit; McMILLAN-MAJOR, Angelina; SHMITCHELL, Shmargaret. **On the dangers of stochastic parrots: can language models be too big?** In: *ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)*. New York: ACM, 2021. p. 610–623.

BOMMASANI, Rishi et al. **On the opportunities and risks of foundation models**. Stanford: Center for Research on Foundation Models, Stanford University, 2021.

CASTELLS, Manuel. **Communication power**. Oxford: Oxford University Press, 2009.

CESARINO, Letícia. **O mundo do avesso: verdade e política na era digital**. São Paulo: Ubu Editora, 2021.

CHESNEY, Robert; CITRON, Danielle Keats. **Deep fakes: a looming challenge for privacy, democracy, and national security**. *California Law Review*, Berkeley, v. 107, n. 6, p. 1753–1820, 2019.

DAVENPORT, Thomas H.; BECK, John C. **The attention economy: understanding the new currency of business**. Boston: Harvard Business School Press, 2001.

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

FLORIDI, Luciano; CHIRIATTI, Massimo. **GPT-3: its nature, scope, limits, and consequences**. *Minds and Machines*, Dordrecht, v. 30, n. 4, p. 681–694, 2020.

FOUCAULT, Michel. **Microfísica do poder**. Organização e tradução de Roberto Machado. Rio de Janeiro: Graal, 1979.

GIDDENS, Anthony. **Modernity and self-identity: self and society in the late modern age**. Cambridge: Polity Press, 1991.

GILLESPIE, Tarleton. **Custodians of the internet: platforms, content moderation, and the hidden decisions that shape social media**. New Haven: Yale University Press, 2018.

HOLMES, Wayne; BIALIK, Maya; FADEL, Charles. **Artificial intelligence in education: promises and implications for teaching and learning**. Boston: Center for Curriculum Redesign, 2019.

LÉVY, Pierre. **Cibercultura**. São Paulo: Editora 34, 1999.

McINTYRE, Lee. **Post-truth**. Cambridge: MIT Press, 2018.

OECD. **Generative AI**. Paris: Organisation for Economic Co-operation and Development, 2024.

OPENAI. **GPT-4 technical report**. arXiv, 2023. DOI: 10.48550/arXiv.2303.08774. Disponível em: <https://arxiv.org/pdf/2303.08774.pdf>. Acesso em: 15 jul. 2026.

PASQUALE, Frank. **The black box society: the secret algorithms that control money and information**. Cambridge: Harvard University Press, 2015.

ROSA, Hartmut. **Social acceleration: a new theory of modernity**. New York: Columbia University Press, 2013.

RUSSELL, Stuart; NORVIG, Peter. **Artificial intelligence: a modern approach**. 4. ed. Hoboken: Pearson, 2021.

SIMON, Herbert A. **Designing organizations for an information-rich world**. In: GREENBERGER, Martin (ed.). *Computers, Communication, and the Public Interest*. Baltimore: Johns Hopkins University Press, 1971. p. 37–72.

SRNICEK, Nick. **Platform capitalism**. Cambridge: Polity Press, 2017.

UNESCO. **Recommendation on the ethics of artificial intelligence**. Paris: UNESCO, 2021.

UNESCO. **Guidance for generative ai in education and research**. Paris: UNESCO, 2023.

UNESCO. **Generative AI and the future of education**. Paris: UNESCO, 2024.

VAN DIJCK, José. **The culture of connectivity: a critical history of social media**. Oxford: Oxford University Press, 2013.

VAN DIJCK, José; POELL, Thomas; DE WAAL, Martijn. **The platform society: public values in a connective world**. Oxford: Oxford University Press, 2018.

VOSOUGHI, Soroush; ROY, Deb; ARAL, Sinan. **The spread of true and false news online.** *Science*, Washington, v. 359, n. 6380, p. 1146–1151, 9 mar. 2018. DOI: 10.1126/science.aap9559. Disponível em: Science. Acesso em: 15 jul. 2026.

ZUBOFF, Shoshana. **The age of surveillance capitalism: the fight for a human future at the new frontier of power.** New York: PublicAffairs, 2019.

CONCLUSÃO

A presente obra teve como propósito compreender os diferentes efeitos da pós-verdade produzidos e impulsionados pela convergência entre plataformas digitais e inteligências artificiais generativas. Longe de tratar tais fenômenos como acontecimentos isolados, buscou-se demonstrar que eles constituem dimensões complementares de uma profunda transformação histórica na forma como as sociedades contemporâneas produzem, organizam, distribuem, legitimam e contestam o conhecimento. A investigação partiu da compreensão de que a crise atual da verdade não pode ser adequadamente explicada por interpretações simplificadoras que atribuem responsabilidade exclusiva às tecnologias digitais, às redes sociais, aos indivíduos ou às disputas políticas. Ao contrário, trata-se de um fenômeno complexo, multidimensional e historicamente constituído, cuja compreensão exige a articulação entre diferentes campos do conhecimento, como Filosofia, Sociologia, Comunicação, Ciência da Informação, Educação, Inteligência Artificial, Ciência Política e Estudos de Plataforma.

O primeiro capítulo evidenciou que a pós-verdade não representa o desaparecimento da verdade nem a simples proliferação de mentiras, mas a transformação das condições culturais, sociais e políticas pelas quais determinadas afirmações passam a ser aceitas coletivamente. Demonstrou-se que o fenômeno possui raízes anteriores ao desenvolvimento das redes digitais e está relacionado à crescente centralidade das emoções, das identidades coletivas, dos vínculos de pertencimento e da crise de legitimidade das instituições tradicionalmente responsáveis pela produção e validação do conhecimento. A análise permitiu compreender que a pós-verdade constitui, sobretudo, uma crise dos mecanismos sociais de confiança, na qual fatos objetivos deixam de exercer, por si sós, força suficiente para organizar o debate público.

O segundo capítulo ampliou essa discussão ao reconstruir historicamente o conceito de redes sociais desde suas origens na Sociologia clássica até sua transformação em plataformas digitais contemporâneas. Demonstrou-se que as redes sociais antecedem em muitas décadas o surgimento da internet e constituem formas históricas de organização das relações humanas. A internet e, posteriormente, as plataformas digitais não criaram as redes sociais, mas

reorganizaram profundamente sua escala, velocidade, visibilidade e capacidade de circulação da informação. O percurso analítico permitiu diferenciar conceitualmente rede social, mídia social, site de rede social e plataforma digital, evidenciando que as atuais plataformas deixaram de funcionar apenas como ambientes de interação para constituírem infraestruturas sociotécnicas responsáveis pela organização dos fluxos informacionais e pela mediação crescente das relações sociais.

O terceiro capítulo concentrou-se nas inteligências artificiais generativas, demonstrando que esses sistemas representam uma inflexão significativa na história da inteligência artificial e da produção do conhecimento. Diferentemente das tecnologias computacionais anteriores, voltadas predominantemente para classificação, previsão ou automação de tarefas específicas, os modelos generativos passaram a produzir linguagem natural, imagens, sons, vídeos e diferentes formas de representação simbólica com elevado grau de sofisticação. Evidenciou-se que tais sistemas ampliam extraordinariamente as possibilidades de produção intelectual, científica, educacional e criativa, ao mesmo tempo em que introduzem novos desafios relacionados à autoria, à confiabilidade, à transparência, à responsabilidade algorítmica, aos vieses, à desinformação e à governança da inteligência artificial.

O quarto capítulo constituiu o núcleo interpretativo desta investigação ao propor que a compreensão da pós-verdade contemporânea exige analisar conjuntamente esses três processos históricos. Defendeu-se que plataformas digitais e inteligências artificiais generativas não apenas ampliam a circulação da informação, mas participam diretamente da reorganização das condições sociotécnicas pelas quais a verdade é produzida, distribuída, validada e contestada. A principal hipótese desenvolvida sustenta que a pós-verdade contemporânea emerge da interação permanente entre dimensões cognitivas, tecnológicas, econômicas, políticas e culturais, configurando um regime algorítmico de mediação do conhecimento no qual sujeitos humanos e sistemas computacionais passam a compartilhar responsabilidades na organização da experiência informacional.

Nesse sentido, uma das principais contribuições desta obra consiste em deslocar o debate da falsa oposição entre determinismo tecnológico e determinismo social. Nem as tecnologias, isoladamente, produzem a pós-verdade, nem

os fenômenos sociais podem ser compreendidos independentemente das infraestruturas técnicas que condicionam a circulação contemporânea da informação. Plataformas digitais, algoritmos, economias baseadas na atenção, inteligências artificiais generativas, instituições científicas, processos culturais e práticas comunicacionais constituem componentes interdependentes de um mesmo ecossistema sociotécnico. A compreensão desse sistema integrado permite superar interpretações reducionistas e oferecer uma leitura mais abrangente dos desafios informacionais do século XXI.

Do ponto de vista teórico, esta pesquisa propõe compreender a pós-verdade não apenas como um fenômeno discursivo, mas como uma propriedade emergente das arquiteturas sociotécnicas contemporâneas. A crise da verdade desloca-se do plano exclusivamente epistemológico para incorporar dimensões infraestruturais relacionadas à organização algorítmica da informação, à platformização da comunicação e à crescente automatização da produção simbólica. Essa perspectiva amplia o escopo tradicional dos estudos sobre pós-verdade ao integrar contribuições provenientes da Filosofia da Informação, dos Estudos de Plataforma, da Sociologia das Redes, da Inteligência Artificial e da Ciência da Informação, propondo um modelo interpretativo interdisciplinar para compreender os diferentes efeitos da convergência entre esses fenômenos.

Do ponto de vista metodológico, o estudo caracteriza-se como um ensaio investigativo fundamentado em ampla revisão crítica da literatura nacional e internacional. Essa opção permitiu articular diferentes tradições teóricas, identificar aproximações conceituais e construir uma interpretação integradora da produção científica recente. Entretanto, essa mesma característica constitui uma de suas principais limitações. Por tratar-se de uma investigação predominantemente teórica, as hipóteses aqui desenvolvidas não foram submetidas à validação empírica mediante estudos de campo, análises quantitativas, etnografias digitais ou experimentos controlados. As proposições apresentadas devem, portanto, ser compreendidas como um modelo explicativo cuja robustez dependerá de futuras investigações empíricas capazes de testar, ampliar, refinar ou eventualmente reformular suas categorias analíticas.

Outra limitação refere-se ao próprio ritmo de transformação das tecnologias digitais. Plataformas, algoritmos e modelos generativos evoluem em velocidade

significativamente superior àquela tradicionalmente observada na produção científica. Consequentemente, parte das discussões aqui apresentadas poderá demandar atualização contínua à medida que novos modelos de inteligência artificial, novos mecanismos de regulação, novas arquiteturas computacionais e diferentes formas de interação sociotécnica forem sendo desenvolvidos. Essa característica não fragiliza a investigação; ao contrário, evidencia que a compreensão da inteligência artificial e da pós-verdade exige permanente revisão conceitual diante da rapidez das mudanças tecnológicas.

Os resultados desta pesquisa também apontam importantes agendas para investigações futuras. A primeira refere-se à necessidade de estudos empíricos capazes de examinar como plataformas digitais e inteligências artificiais generativas influenciam concretamente processos de aprendizagem, formação da opinião pública, produção científica, práticas jornalísticas e construção da confiança social. A segunda envolve o desenvolvimento de modelos teóricos e metodológicos destinados à avaliação da transparência, da explicabilidade e da governança de sistemas algorítmicos utilizados em diferentes contextos sociais. Uma terceira agenda diz respeito às implicações educacionais desse novo cenário, especialmente no que se refere à alfabetização midiática, informacional, algorítmica e em inteligência artificial como competências fundamentais para a cidadania contemporânea. Finalmente, torna-se imprescindível ampliar as pesquisas sobre ética, regulação e direitos digitais, buscando construir modelos de governança capazes de conciliar inovação tecnológica, proteção dos direitos fundamentais, pluralidade informacional e fortalecimento democrático.

As reflexões desenvolvidas ao longo deste livro conduzem, por fim, a uma compreensão que ultrapassa o debate específico sobre plataformas digitais ou inteligência artificial. A principal questão contemporânea talvez não seja simplesmente determinar se determinada informação é verdadeira ou falsa, mas compreender quais condições sociais, políticas, tecnológicas e culturais tornam possível o reconhecimento coletivo da verdade em sociedades profundamente mediadas por sistemas computacionais. Em um contexto caracterizado pela crescente automatização da produção simbólica, pela plataformização da comunicação e pela expansão das inteligências artificiais generativas, preservar a verdade deixa de significar a defesa de uma autoridade infalível e passa a representar o

fortalecimento permanente de procedimentos públicos de investigação, crítica, transparência, responsabilidade e diálogo.

A verdade, compreendida nessa perspectiva, não constitui um patrimônio estático nem uma imposição institucional incontestável. Ela permanece sendo uma construção social continuamente submetida ao exame crítico, ao confronto de evidências, à revisão metodológica e ao debate democrático. O desafio do século XXI não consiste em restaurar modelos idealizados de autoridade epistemológica, tampouco em aceitar o relativismo que dissolve toda possibilidade de conhecimento compartilhado. O verdadeiro desafio consiste em construir ecossistemas sociotécnicos nos quais inovação tecnológica, inteligência artificial, liberdade de expressão, produção científica e responsabilidade pública possam coexistir sem comprometer as condições que tornam possível a própria vida democrática.

Conclui-se, portanto, que os diferentes efeitos da pós-verdade gerados e impulsionados pelas redes sociais e pelas novas inteligências artificiais generativas não devem ser interpretados apenas como ameaças decorrentes do avanço tecnológico. Eles expressam uma transformação histórica muito mais profunda: a reorganização das formas pelas quais a humanidade produz, compartilha, interpreta e legitima o conhecimento. Compreender essa transformação representa um dos grandes desafios científicos, educacionais, políticos e éticos do nosso tempo. Enfrentá-la exigirá não apenas tecnologias mais transparentes ou regulações mais eficientes, mas, sobretudo, uma renovada capacidade coletiva de construir confiança, responsabilidade e compromisso com a produção social do conhecimento em uma sociedade crescentemente conectada, algorítmica e inteligente.

Longe de encerrar o debate, esta investigação procura oferecer um referencial teórico para compreender uma realidade em permanente transformação. Nesse horizonte, permanecem abertas inúmeras possibilidades de pesquisa, entre as quais destacam-se as seguintes questões:

- 1) Se plataformas digitais e inteligências artificiais generativas passaram a integrar as infraestruturas responsáveis pela mediação da verdade, quais transformações ainda poderão ocorrer nos regimes contemporâneos de produção, legitimação e circulação do conhecimento diante da crescente autonomia dos sistemas algorítmicos?

2) De que maneira diferentes contextos sociais, políticos, culturais e educacionais modificam os efeitos da convergência entre pós-verdade, plataformas digitais e inteligências artificiais generativas na construção da confiança pública e da autoridade epistemológica?

3) Quais estratégias educacionais, institucionais e regulatórias poderão fortalecer processos democráticos de validação do conhecimento sem comprometer a pluralidade de perspectivas, a liberdade de expressão e a inovação tecnológica diante dos novos regimes da verdade?

A verdade nunca foi uma construção concluída. Em cada época, ela é permanentemente produzida, disputada e reinterpretada. Compreender os novos regimes da verdade significa, antes de tudo, compreender as transformações da própria sociedade contemporânea — um desafio que permanece aberto às próximas gerações de pesquisadores.

MARCELO JOSÉ HANAUER

OS NOVOS REGIMES DA VERDADE



PÓS-VERDADE, PLATAFORMAS DIGITAIS
E INTELIGÊNCIA ARTIFICIAL

COMO A CONVERGÊNCIA ENTRE PÓS-VERDADE,
PLATAFORMAS DIGITAIS E INTELIGÊNCIAS ARTIFICIAIS
REDEFINE A PRODUÇÃO, A CIRCULAÇÃO E A LEGITIMAÇÃO
DO CONHECIMENTO NA SOCIEDADE CONTEMPORÂNEA