

Desafios éticos e sociotécnicos associados ao uso de algoritmos em redes sociais

Ethical and Sociotechnical Challenges of Algorithm Use in Social Networks

Regina Rossetti

<https://orcid.org/0000-0002-3219-4570>

Doutora em Filosofia. Universidade de São Paulo (USP) – Brasil.
regina.rossetti@online.uscs.edu.br

Kethly Garcia

<https://orcid.org/0009-0008-0092-3338>

Mestre em Comunicação. Universidade São Caetano do Sul (USCS) – Brasil. kethly.garcia@online.uscs.edu.br

Magda Aparecida Silveiro Miyashiro

<https://orcid.org/0009-0007-0322-6143>

Doutora em Engenharia e Tecnologia Espaciais. Instituto de Nacional Pesquisas Espaciais (INPE) – Brasil.
magda.miyashiro@online.uscs.edu.br

RESUMO

Os sistemas de informação, que envolvem o uso de inteligência artificial (IA) e algoritmos, tornaram-se indispensáveis para grande parte das empresas, incluindo as plataformas responsáveis pela gestão e pelo controle das redes sociais *on-line*. Esses sistemas algorítmicos presentes nas redes sociais *on-line* apresentam desafios aos gestores e riscos aos usuários. O presente artigo teve por objetivo identificar e discutir, com base em uma pesquisa bibliográfica, os principais desafios éticos e sociotécnicos associados ao uso de algoritmos em redes sociais *on-line*, por meio de abordagem qualitativa. Os dados foram analisados mediante categorização temática, o que possibilitou a identificação de dez desafios centrais. Os resultados indicam que o surgimento de sistemas técnicos que utilizam algoritmos na gestão e mediação das relações e ações humanas nas redes sociais *on-line* apresenta desafios relacionados à modulação, à arrogância epistêmica, à opacidade, aos interesses privados, ao capitalismo de vigilância, à mineração de dados, à falibilidade, aos vieses, à segregação e à pós-verdade.

Palavras-chave: Redes sociais; Algoritmos; Inteligência Artificial; Gestão tecnológica.

ABSTRACT

Information systems, which involve the use of artificial intelligence (AI) and algorithms, have become indispensable for a large share of companies, including platforms responsible for the management and control of online social networks. These algorithmic systems embedded in online social networks pose challenges for managers and risks for users. This article aimed to identify and synthesize, through a qualitative approach, the main challenges associated with these systems, based on a bibliographic review. The data were analyzed through thematic categorization, enabling the identification of ten central challenges. The results indicate that the emergence of technical systems that employ algorithms in the management and mediation of human relationships and actions in online social networks presents challenges related to modulation, epistemic arrogance, opacity,

private interests, surveillance capitalism, data mining, fallibility, bias, segregation, and post-truth.

Keywords: social networks; Algorithms; Artificial Intelligence; Technological management.

Recebido em 31/05/2026. Aprovado em 16/03/2026. Avaliado pelo sistema *double blind peer review*. Publicado conforme normas da ABNT.
<https://doi.org/10.22279/navus.v18.2147>

1 INTRODUÇÃO

As redes sociais on-line, a princípio, se caracterizam como sendo estruturas do ambiente digital que têm por objetivo promover interações entre seus participantes, por meio do compartilhamento de interesses em comum. Entretanto, com o tempo, passaram a apresentar outras funcionalidades, principalmente, monetárias, por serem associadas às interações com as empresas que vendem produtos e serviços e com as agências que fornecem notícias e informações. Essa apropriação das redes sociais pela gestão de negócios ocorreu em função da abrangência e alcance dessas redes em direção aos públicos de interesse. Conforme o site *We Are Social* (2022), em parceria com a *Hootsuite*, responsável por relatórios diversos relacionados ao uso da internet e de mídias sociais, redes sociais e plataformas de serviços diversas ao redor do mundo, o número de utilizadores de redes sociais on-line triplicou nas últimas décadas e supera o número de utilizadores da internet do mesmo período, enquanto o tempo de uso das redes sociais on-line durante o dia vem aumentando exponencialmente a cada ano, o que representa sua predominância em nossas vidas.

A forma como o ser humano contemporâneo se relaciona e os processos comunicacionais que o envolvem estão, mais do que nunca, envoltos por esses ambientes digitais, que, por sua vez, contam com sistemas técnicos que envolvem algoritmos e inteligência artificial para intermediar essas relações. Os algoritmos e sistemas técnicos, embora sejam necessários para gerir a imensa quantidade de dados produzida diariamente pelas redes, apresentam também muitos desafios. Esses desafios algorítmicos são citados por diversos autores ao longo dos anos, e, por meio de uma revisão bibliográfica, foram identificados os principais desafios algorítmicos e sua relação como intermediários na comunicação neste estudo.

Este estudo tem como objetivo identificar e discutir, a partir da literatura científica, os principais desafios éticos e sociotécnicos associados ao uso de algoritmos em redes sociais on-line. Nesse contexto, busca-se responder à seguinte questão de pesquisa: quais são os principais desafios éticos e sociotécnicos associados ao uso de algoritmos em redes sociais on-line?

2 FUNDAMENTAÇÃO TEÓRICA

A seguir, apresentam-se os conceitos fundamentais que sustentam a análise dos desafios algorítmicos, organizados em torno da teoria da comunicação, da ética algorítmica e da gestão tecnológica contemporânea.

2.1 Teoria da Comunicação e o Agir Comunicativo

A comunicação é compreendida como um dos atos fundamentais para o desenvolvimento do ser (Gomes, 2006). Ela se caracteriza pelo que é comum, compartilhado, geral e público, envolvendo a partilha de informações e a coparticipação no espaço em que essa partilha ocorre (Gomes, 2006). Segundo a teoria de Habermas (1981), a ação comunicativa ocorre plenamente no mundo vivido, sendo pautada em princípios de solidariedade, cooperação e transparência.

Contudo, no mundo sistêmico, o poder assume a forma de ação instrumental, passando a colonizar o mundo vivido e a infiltrar-se nas

relações sociais (Habermas, 1981). Quando a instrumentalidade e as formas de poder se sobrepõem à ética discursiva, a comunicação deixa de ser legítima e justa. Nesse contexto, Matellart (1999) defende a necessidade de reconfiguração das funções da comunicação, de modo a superar o tecnicismo e as lógicas globais mercantis, com foco nas demandas vinculadas ao interesse público.

2.2 Ética Algorítmica e Sociedade de Controle

Os algoritmos não são recursos neutros, pois resultam da tradução humana de posicionamentos, valores e ideologias para a linguagem computacional (Santana, 2022). Esse caráter frequentemente obscuro remete ao conceito de "caixa preta", proveniente da cibernética de Flusser e da eletrônica de Bateson, que caracteriza os algoritmos como recursos inescrutáveis e inacessíveis aos usuários (Santana, 2022).

Sob a lógica do capitalismo de vigilância, conforme Shoshana Zuboff (2019), emerge o conceito de poder instrumentário, que consiste na análise e na modulação do comportamento humano em prol de finalidades de terceiros e geração de receita. Como afirma Foucault (1996), o exercício do poder hierárquico sobre um objeto requer, primeiramente, observá-lo e conhecê-lo. No ambiente digital, a utilização abusiva de algoritmos para monitoramento sugere a constituição de uma sociedade controlada, o que se opõe ao empoderamento dos usuários (Martino; Marques, 2014). Além disso, preconceitos estruturais como racismo e sexismo, podem ficar impressos nesses sistemas (Noble, 2018).

2.3 Gestão Tecnológica e Sistemas de Informação

A gestão contemporânea é definida como uma abordagem adaptada aos desafios do ambiente de negócios moderno, caracterizado por mudanças rápidas, avanços tecnológicos e elevado nível de interconectividade entre as organizações. Os sistemas de informação que integram inteligência artificial (IA) e algoritmos tornaram-se ferramentas indispensáveis para a gestão e o controle das redes sociais on-line.

Nesse cenário, os algoritmos atuam na estruturação silenciosa das práticas cotidianas e na mediação dos processos comunicacionais (Fernández, 2021; Ramírez; Jiménez, 2020). O desafio para a gestão reside no fato de que tais sistemas, embora úteis, apresentam falibilidade e imprevisibilidade, podendo violar direitos pessoais quando a automatização ocorre sem intervenção humana adequada (Mittelstadt, 2016; Pant, 2023).

3 MATERIAIS E MÉTODOS

Este estudo caracteriza-se como uma pesquisa bibliográfica de abordagem qualitativa, desenvolvida por meio de uma revisão narrativa, analisando 21 estudos publicados entre 2010 e 2023, provenientes do Portal CAPES, BRAPCI, SSRN, arXiv, entre outros.

Para garantir a coerência do estudo, a estratégia PICO (População, Intervenção, Comparação/Contexto e Outcome) foi utilizada como guia não apenas para a formulação da pergunta de pesquisa, mas também para a definição do título e do objetivo central. A aplicação do acrônimo permitiu que o título fosse definido a partir dos componentes População (redes sociais on-line) e Intervenção (algoritmos inteligentes). Da mesma forma, o

objetivo foi estruturado de modo a explicitar o Desfecho (Outcome) esperado, qual seja, a identificação e a discussão ética dos desafios sociotécnicos.

Embora o acrônimo PICO seja tradicionalmente empregado em revisões sistemáticas, especialmente na área da saúde, neste estudo ele foi utilizado de forma adaptada como ferramenta heurística de organização conceitual. Seu uso não teve como finalidade estruturar um protocolo rígido de seleção e análise, mas sim orientar a delimitação do problema de pesquisa, a definição do escopo e a construção do objetivo do estudo, contribuindo para a coerência interna da investigação.

Essa abordagem foi utilizada para assegurar que o escopo da revisão bibliográfica permanecesse estritamente alinhado ao problema de pesquisa proposto, conforme apresentado no Quadro 1.

Quadro 1 - Utilização do Acrônimo PICO

Letra	Significado	Aplicação na pergunta	Termos de busca	Objetivo analítico
P	População / Problema – quem ou o que está sendo estudado?	Redes sociais online	rede social; mídia social; plataforma	Delimitar o contexto do estudo
I	Intervenção – o que está sendo feito ou analisado?	Uso de algoritmos	algoritmo	Definir o fenômeno analisado
C	Comparação adaptado para Contexto – Qual é o cenário do estudo?	Inteligência artificial	inteligência artificial	Caracterizar o ambiente tecnológico
O	Resultados (Outcomes) – o que se quer observar, entender ou medir?	Desafios éticos e sociotécnicos associados ao uso de algoritmos	ética; princípio ético	Identificar e organizar os desafios

Foram selecionados estudos que abordassem a concepção e a atuação de algoritmos em redes sociais on-line, através das palavras-chave: inteligência artificial, ética, princípio ético, rede social, mídia social, plataforma, algoritmo e suas possíveis variações no plural. A partir dessa busca, foram identificados 23 artigos, dos quais 21 foram selecionados e dois foram excluídos por não abordarem, de forma relevante, problemas relacionados à concepção e à atuação de algoritmos em redes sociais.

Para a análise dos dados, foi realizada a categorização temática dos 21 artigos selecionados. Foram identificados diversos termos correlatos que apresentavam sentidos similares na literatura. Esses termos foram agrupados em 10 categorias centrais, denominadas 'desafios'. Com base nessa categorização, procedeu-se à síntese crítica dos principais problemas éticos e sociotécnicos enfrentados pela gestão tecnológica em redes sociais.

Dessa forma, a análise considerou os 21 artigos selecionados como base empírica, a partir dos quais foram identificados 10 desafios principais abordados pelos autores, bem como os conceitos correlatos a eles associados.

Ressalta-se que ferramentas de inteligência artificial generativa foram utilizadas como apoio na revisão linguística e na organização textual, sem interferência na interpretação dos dados ou na elaboração das análises.

4 RESULTADOS E DISCUSSÃO

Esta seção apresenta e discute os resultados obtidos por meio da análise temática, os quais correspondem à identificação de 10 desafios centrais relacionados à gestão e à mediação de algoritmos em redes sociais, a saber: modulação, arrogância epistêmica, opacidade, interesses privados, capitalismo de vigilância, mineração de dados (*data mining*), falibilidade, viés, segregação e pós-verdade. Durante a análise, observou-se que diferentes autores utilizam termos correlatos para tratar conceitos com significados similares; por essa razão, esses termos foram sistematicamente agrupados por afinidade semântica em categorias temáticas, que consolidam discussões de mesmo sentido, garantindo maior rigor analítico e sistematização crítica dos achados. A síntese desses resultados, incluindo suas descrições, conceitos associados e impactos éticos, é apresentada no Quadro 2, servindo de base para o detalhamento individual de cada desafio.

Quadro 2 - Síntese dos desafios éticos e sociotécnicos identificados na literatura

Síntese dos desafios				
Desafio	Descrição	Conceitos Correlatos	Impacto Ético/Social	Autores Citados
Modulação	Processo de influência nas opiniões e ações do agente em quatro etapas: identificação, perfilamento, acompanhamento e atuação.	Profiling, antecipação do usuário, filtragem de conteúdo, algoritmos de relevância e personalização, manipulação.	Redução da autonomia na construção de crenças e interferência negativa nas ações comunicativas.	Silveira (2019), Santana (2022), Vries (2010), Gillespie (2018), Dijck (2013), Mittelstadt (2016), Fernández (2021), Moraes e Testa (2020).
Arrogância Epistêmica	Fenômeno onde o agente assume saber mais do que realmente sabe devido à confirmação constante do ambiente digital.	Filtros-bolha (bolhas epistêmicas), câmaras de eco e polarização.	Retração do processo comunicativo e interferência direta em desfechos democráticos e na formação da opinião pública.	Moraes e Testa (2020), Pedrosa e Junior (2021), Santaella (2018), Santana (2022), Stroppa (2021).
Opacidade	Inescrutabilidade dos caminhos percorridos pelos algoritmos, tornando seus códigos e decisões inacessíveis.	Algoritmos de caixa preta e inescrutabilidade.	Dificulta a auditoria e o controle democrático, pois as regras são tratadas como segredos comerciais.	Santana (2022), Pedrosa e Junior (2021), Silva, Aragão e Pretto (2021), Silveira (2019), Mittelstadt (2016), Ramirez e Jiménez (2019).
Interesses Privados	Atuação estratégica voltada para atingir interesses individuais e maximização de lucro das plataformas.	Uso de dados para marketing e marketing predatório baseado na vulnerabilidade do usuário.	Sobreposição de interesses privados ao interesse público e distanciamento de poder entre detentores do código e o cidadão.	Silveira (2019), Caribé (2017), Matellart (1999).
Capitalismo de Vigilância	Lógica de poder que analisa, adapta e remodela o comportamento humano para gerar receita para terceiros.	Invasão à privacidade, monitoramento constante e risco à proteção de dados.	Estabelecimento de uma sociedade controlada e colonização do "mundo vivido" pelo "mundo sistêmico".	Zuboff (2019), Fernández (2021), Ferreira, Pinheiro e Marques (2021), Garcia (2021), Martino e Marques (2014), Moraes e Testa (2020), Foucault (1996), Habermas (1981).
Mineração de Dados	Etapa fundamental de extração e coleta de dados massivos para processamento algorítmico.	Coleta de dados, termos de uso/serviço e dependência tecnológica.	Vulnerabilidade dos usuários às condições impostas e risco de falhas de processamento em correlações de sentido.	Moraes e Testa (2020).
Falibilidade	Incapacidade dos sistemas automatizados de serem infalíveis, apresentando erros em classificações e decisões.	Imprevisibilidade, erros de processamento, volatilidade e auto-tagging.	Violação de direitos pessoais e perda de controle do desenvolvimento da ferramenta pelos próprios programadores.	Caribé (2017), Gillespie (2018b), Pant (2023), Mittelstadt (2016).
Viés	Natureza não neutra dos algoritmos, que carregam posicionamentos e ideologias de seus criadores.	Ausência de neutralidade e falta de imparcialidade.	Reprodução de racismo, sexismo e visões estereotipadas devido à falta de diversidade no mercado tecnológico.	Santana (2022), Noble (2018).
Segregação	Tratamento diferenciado entre usuários que perpetua a discriminação e desigualdades na rede.	Discriminação, preconceito e desigualdades estruturais.	Violação do direito ao tratamento igualitário e invisibilização de temas importantes para o desenvolvimento social.	Fernández (2021), Mittelstadt (2016), Pant (2023).
Pós-Verdade	Uso de algoritmos para a produção e disseminação em massa de conteúdos falsos ou desinformação.	Fake news, deep fakes, golpes, falsificações e desinformação.	Confusão na percepção da realidade pelo público e potencial de produzir consequências sociais graves.	Moraes e Testa (2020), Garcia (2021).

Observa-se que os desafios identificados não se apresentam de forma isolada, mas como elementos interdependentes, compondo um conjunto de implicações sociotécnicas que se articulam e se reforçam mutuamente.

4.1 Modulação

Os conceitos correlatos à modulação algorítmica incluem formação de perfil, *profiling*, antecipação do usuário, filtragem de conteúdo, algoritmos de relevância pública, Identificação de relevância, algoritmos de personalização, influencia algorítmica, influência política, controle e manipulação algorítmicos.

O conceito de modulação algorítmica, ou simplesmente modulação, conforme Silveira (2019), envolve quatro etapas principais, apresentadas na seguinte ordem: a identificação do agente, a formação de perfil do agente, a construção de processos de acompanhamento do agente e, por fim, a atuação sobre o agente identificado, a fim de modular seu comportamento. Santana (2022) também aborda esse conceito, exemplificando a completude e abrangência das suas quatro etapas.

Na etapa de identificação do agente, realiza-se a coleta de dados relativos ao usuário da rede, sendo essa coleta geralmente realizada mediante consentimento, seja por meio do aceite de cookies e termos de serviço das plataformas, seja de forma indireta, pelo aceite de terceiros a termos de serviços de determinadas empresas que podem coletar dados de contatos¹. Essa coleta, inclusive, envolve discussões importantes sobre dados, privacidade e autodeterminação do usuário.

Já a segunda etapa do referido processo consiste na formação de perfil, ou *profiling* (Vries, 2010) bem como na antecipação do comportamento do usuário (Gillespie, 2018; Vries, 2010). Ao antecipar o usuário utilizando a técnica de *profiling* ou formação de perfil, são realizadas suposições e induções, constituindo uma espécie de "persona" baseada em análises estatísticas que, por se fundamentarem em probabilidades, nem sempre consideram a racionalidade e as subjetividades dos usuários, podendo, inclusive, contribuir para o reforço de estereótipos (Vries, 2010), os quais se perpetuam nas redes sociais on-line e nos sistemas técnicos. A antecipação do usuário configura-se como uma tentativa, por parte dos provedores de algoritmos, de delimitar características comportamentais, e, conforme Gillespie (2018), a dissociação entre probabilidade e realidade factual pode gerar tanto benefícios quanto malefícios. Vries (2010) destaca ainda que essa antecipação da identidade suscita importantes questões filosóficas e éticas, uma vez que, na contemporaneidade, a identidade é composta e estabelecida não apenas a partir da relação, interação e construção com seres humanos, mas também por tecnologias diversas.

Em relação à construção do processo de acompanhamento, terceira etapa do processo de modulação, trata-se, por definição, de um conjunto de dispositivos e técnicas responsáveis por analisar as ações dos agentes, incluindo também os dispositivos e técnicas infiltrados no cotidiano (Silveira, 2019), seja por desconhecimento do usuário ou pela proteção das informações por parte das plataformas que gerenciam redes sociais on-line, frequentemente orientadas por interesses privados.

¹ Ao criar uma conta no Facebook, um dos documentos disponíveis na tela de aceite chama-se "Informações para pessoas que não usam os Produtos da Meta", documento que trata da coleta de dados de contatos dos usuários, independente de esses terem aceito fazer parte da rede ou não.

Na atuação sobre o agente, quarta etapa do processo de modulação, ocorre uma influência em relação as opiniões e ação do agente modulável. Nessa etapa, podem ser inseridos diversos conceitos correlatos, como: filtragem de conteúdo, algoritmos de relevância pública, identificação de relevância, algoritmos de personalização, influência algorítmica, influência política, controle e manipulação algorítmica.

As plataformas, embora se apoiem no discurso de neutralidade e imparcialidade, contam com recursos que filtram conteúdos e definem seu nível de exposição na rede (Dijck, 2013). Karlsson, Van Couvering e Lindell (2022) afirmam que essa atribuição e definição de valor em relação aos conteúdos estabelece certas lógicas de controle que produzem consequências sociais, políticas e culturais relevantes. Desse modo, verifica-se a existência de uma intermediação da informação através da filtragem de conteúdo, a qual contribui para a modulação do comportamento social.

Gillespie (2018), ao citar os algoritmos de relevância pública e a identificação de relevância, afirma que a definição e o direcionamento do que é ou não relevante para um usuário é um ato carregado de significado, e devido à complexidade desse julgamento, muitas vezes ocorre um direcionamento de conteúdos que geram satisfação, o que torna o conceito de relevância evasivo e superficial (Gillespie, 2018).

Para Mittelstadt (2016), existe uma linha tênue entre apoiar decisões dos usuários e influenciá-las ou controlá-las por meio dos algoritmos de personalização de conteúdo. Ao direcionar determinados conteúdos para os usuários, um algoritmo pode tanto auxiliar a tomada de decisão quanto induzi-la ou persuadi-la.

Os algoritmos atuam, portanto, sobre a estruturação silenciosa que é feita na vida das pessoas, assim como, influentes sobre o delineamento de práticas cotidianas. Nesse sentido, evidencia-se a presença de uma influência algorítmica atuante sobre a vivência dos indivíduos (Fernández, 2021; Ramírez; Jiménez, 2020). Moraes e Testa (2020), por sua vez, citam a manipulação e controle sobre as atividades dos usuários, questionando, inclusive, as garantias de autonomia na construção de crenças e opiniões por parte dos indivíduos atuantes na rede.

Dadas as correlações de sentidos entre os conceitos abordados, propõe-se a inclusão de todos os termos mencionados como parte do constructo de modulação algorítmica. A modulação é um conceito que demonstra como os algoritmos são artefatos responsáveis por intermediar os processos comunicacionais nas redes sociais on-line. Dessa forma, a modulação algorítmica interfere diretamente no desenvolvimento do ser e nos princípios de transparência e cooperação discutidos na fundamentação teórica.

4.2 Arrogância epistêmica

Outro desafio identificado na análise temática refere-se ao conceito de arrogância epistêmica, que se relaciona semanticamente com termos como filtros-bolha, câmaras de eco e polarização.

Embora alguns autores utilizem esses conceitos como sinônimos, para Moraes e Testa (2020), eles podem ser diferenciados. Conforme os autores, as câmaras de eco se referem aos espaços em que a opinião e a visão do usuário são ecoadas, ou seja, os algoritmos direcionam conteúdos afins aos gostos do usuário com frequência, atuando com base em um reforço da personalidade e do individualismo. A partir desse processo, os usuários

passam a ter a sensação de que seu ponto de vista é validado pelo mundo, ainda que isso não configure a realidade. Uma das características da câmara de eco é que, apesar desse reforço de gostos e identidade, não se exime a possibilidade de ter contato com outros pontos de vista, embora isso ocorra com baixa frequência.

Em contrapartida, os filtros-bolha, também conhecidos como bolhas epistêmicas (Pedrosa; Junior, 2021), ou *filter bubble*, conceito inicialmente cunhado por Eli Pariser em obra homônima (Santaella, 2018), referem-se a ambientes configurados por algoritmos que restringem o acesso do usuário a qualquer informação conflitante com seu ponto de vista, isolando-os nas chamadas bolhas ideológicas que passam uma falsa percepção sobre o mundo e possuem influência na polarização. Esses dois fenômenos, portanto, convergem para a constituição da arrogância epistêmica, caracterizada como a situação em que um agente assume ter mais conhecimento e saber mais do que sabe na prática, por meio de uma confirmação que se dá no ambiente que o cerca.

Esses fenômenos são invisíveis na rede, embutem-se nas práticas cotidianas, e conforme Santana (2022), tornando-se impraticável entender as suas delimitações e suposições, devido à opacidade algorítmica. Uma das razões de ser da comunicação, que integra seus significados fundamentais, é o que é comum (a todos ou a muitos), compartilhado, geral e público. Comunicar consiste em compartilhar informações e coparticipar do espaço em que essa partilha ocorre, podendo manifestar-se de diversas formas, por meio de múltiplas linguagens e em diferentes esferas (Gomes, 2006).

A arrogância epistêmica configura-se como um processo que influencia as ações comunicacionais, na medida em que restringe, em vez de expandir, as interações em rede e o contato com realidades diversas, exercendo influência direta na formação de opinião. Para Stroppa (2021), a formação da opinião decorre da manifestação de desejos, os quais se transformam em demandas que contribuem para a constituição do interesse público. Uma vez que a esfera pública tem a influência do uso de tecnologias diversas, como a inteligência artificial e os algoritmos que mediam tais manifestações, verifica-se a existência de uma interferência direta nos processos de formação da opinião pública, com implicações relevantes para os desfechos democráticos.

4.3 Opacidade

No âmbito da opacidade algorítmica, identificam-se como conceitos correlatos os algoritmos de caixa-preta (ou simplesmente caixa-preta) e a inescrutabilidade.

Um dos principais desafios relacionados aos algoritmos refere-se à opacidade. Esse conceito também estabelece relações com o conceito de algoritmos de caixa preta² (ou apenas caixa preta) e inescrutabilidade. O conceito caixa preta origina-se na cibernética de Flusser e na eletrônica de Bateson, remetendo aos códigos inacessíveis desses recursos (Santana, 2022). Conforme Pedrosa e Junior (2021), Silva, Aragão e Pretto (2021) e Silveira (2019), os algoritmos são recursos tecnológicos obscuros e se

² Dá-se o nome de caixa preta a um dispositivo fechado, ao qual não se tem acesso completo e do qual apenas seria possível extrair suposições de funcionamento com base em seus *inputs* (entradas) e *outputs* (saídas). Bateson aplica o conceito para questões científicas e filosóficas, no sentido de nomear os problemas incompreensíveis racionalmente e apenas intuídos como as caixas pretas. O encontro da caixa preta de Flusser e de Bateson se dá no ponto de significado de algo desconhecido (Machado, 2005).

utilizam de mecanismos opacos frente aos usuários do ciberespaço. Considerando que os algoritmos atuam no processamento de dados e na tomada de decisões, conforme Mittelstadt (2016), seria adequado supor que os caminhos percorridos durante sua atuação fossem acessíveis. No entanto, na maioria dos casos, isso não se verifica, devido à inescrutabilidade dos algoritmos.

Os algoritmos são, portanto, pouco explicáveis e de difícil controle, sendo raramente transparentes, explicáveis ou auditáveis. Para Ramírez e Jiménez (2019), as grandes corporações de tecnologia, as famosas Big Techs, não delinham claramente suas regras e instruções de uso dos algoritmos, uma vez que muitos desses sistemas são protegidos como segredos comerciais, garantindo vantagens competitivas no mercado.

4.4 Interesses privados

No contexto dos interesses privados que orientam a atuação das plataformas digitais, destacam-se como conceitos correlatos o uso de dados para fins de marketing e o marketing predatório.

Ao falar sobre recursos tecnológicos que são propriedade de empresas privadas, é esperado que atuem de forma estrategicamente planejada, a fim de atingir interesses individuais. O êxito de plataformas de serviços, de mídias e redes sociais ao longo do tempo fez com que houvesse um gradativo aumento dos modelos de negócios que envolvem os sujeitos que oferecem serviços e aqueles que demandam serviços e mercadorias nesses espaços.

Com o início do armazenamento sistemático de dados dos usuários para fins de análise por empresas de marketing, essas plataformas passaram a direcionar conteúdos e ofertas com maior precisão, obtendo resultados mais eficazes. Esse processo ampliou significativamente o poder dessas organizações, uma vez que, por meio de bases de dados estruturadas, tornou-se possível realizar cruzamentos informacionais para a oferta segmentada de publicidade e serviços. Ao longo do tempo, os algoritmos consolidaram-se como ferramentas centrais para a maximização de lucros no contexto do marketing digital, reforçando a lógica econômica que orienta a atuação dessas plataformas.

Diferentemente das mídias tradicionais, que se baseiam na veiculação de conteúdos publicitários direcionados a uma audiência ampla e relativamente homogênea, as plataformas digitais adotam estratégias de microtargeting³, mais direcionadas e ainda conforme aborda Silveira (2019), mais eficiente. Para Caribé (2017) os algoritmos usados para gerir dados e desenhar padrões que auxiliarão empresas a lançar ações publicitárias, alimentam o chamado marketing predatório, baseado na vulnerabilidade dos usuários. Essa ação traz resultados positivos para o anunciante e demonstra a distância de poder entre os cardeais do algoritmo (ou seja, os que detêm conhecimento e domínio em relação a esses recursos) e o cidadão comum. Esses interesses privados levam as plataformas a atuarem de maneira que os interesses privados se sobreponham ao interesse público, rompendo com princípios que envolvem a coletividade.

Para Matellart (1999), é necessário, mais do que nunca, introduzir novas funções de comunicação na sociedade, que não a compreendam apenas como um recurso subordinado às lógicas globais mercantis. A comunicação,

³ Trata-se do método de fazer previsões personalizadas para entender os usuários, seu hábitos, costumes e gostos e dessa forma direcionar propagandas específicas. Além do direcionamento de produtos e serviços, a técnica pode ser usada para influência política e outros fins (Freire, 2020).

diferentemente do que tem sido tradicionalmente ensinado, não deve assumir um caráter estritamente tecnicista nem formar profissionais orientados exclusivamente por essa perspectiva, mas sim voltar-se às demandas que ultrapassam o entretenimento e contemplam as necessidades do interesse público. Para o autor, os meios e os profissionais da comunicação devem pensar no interesse público e não diminuir a representação de certas necessidades, mas sim destacá-las frente aos diversos agentes.

4.5 Capitalismo de vigilância

Os usuários de plataformas estão, a cada dia, sendo mais analisados, alcançados e remodelados. Esse processo está inserido em uma lógica estruturada, que orienta o uso e o direcionamento de recursos tecnológicos em favor de interesses particulares.

Para Zuboff (2019), a essa lógica dá-se o nome de capitalismo de vigilância. Esse capitalismo gera uma nova forma de poder, chamada de instrumentarismo, e o poder instrumentário é o que analisa e adapta o comportamento humano em prol das finalidades de terceiros, que o usarão para gerar receita. O capitalismo de vigilância é um conceito ligado à invasão da privacidade, ao risco à privacidade e à proteção de dados, à visibilidade e ao monitoramento.

Para Fernández (2021), com a utilização abusiva de algoritmos, estamos diante da possibilidade de uma sociedade altamente controlada e também controladora. Ferreira, Pinheiro e Marques (2021) afirmam que o uso de dados para gerar informações que alimentem práticas de vigilância já é uma preocupação da sociedade civil. De forma complementar, Garcia (2021) destaca que é possível o uso indevido de dados gerenciados por algoritmos para afetar cidadãos, de forma não intencional ou deliberada. Essa visibilidade e vigilância de usuários do ciberespaço sugerem uma posição antagônica ao empoderamento ou fortalecimento dos usuários dos ambientes digitais (Martino; Marques, 2014).

Moraes e Testa (2020) exemplificam que essa vigilância se intensifica no contexto da computação ubíqua⁴, em que dispositivos como câmeras de vigilância, por exemplo, não apenas armazenam dados sobre os cidadãos, mas também se articulam a outros sistemas, fornecendo informações relevantes para algoritmos capazes de traçar perfis abrangentes e utilizar esses dados para fins de controle. Nesse sentido, a questão ultrapassa dimensões estritamente comerciais, configurando-se como uma relação de controle e dominação social exercida por atores que detêm poder. Nesse contexto, conforme afirma Foucault (1996), o exercício do poder hierárquico pressupõe, primeiramente, a observação e o conhecimento do objeto, o que reforça o papel da vigilância como mecanismo central de controle.

Os processos comunicacionais, no entanto, para que sejam efetivos e legítimos, devem estar livres de assimetrias de poder que comprometam a interação. Essa lógica de vigilância evidencia a infiltração do "mundo sistêmico" sobre o "mundo vivido" dos usuários, conforme a teoria de Jürgen Habermas, prejudicando a legitimidade da comunicação nas redes sociais.

⁴ Computação ubíqua refere-se à normalização na tecnologia nas práticas cotidianas, tornando a relação humano-máquina cada vez mais comum e até imperceptível, dada a sua naturalização na sociedade (UFPR, 2022).

4.6 Data mining

O conceito de *data mining*, ou mineração de dados, basicamente se relaciona com os outros conceitos apresentados, uma vez que essa é a etapa inicial e fundamental para uma base sólida que se transforme em informação para o processamento dos algoritmos. Esse conceito relaciona-se à coleta de dados, e a discussão por trás do *data mining* está justamente nos algoritmos que são utilizados para a extração de dados sem consentimento dos usuários ou para atingir fins ilegítimos.

As empresas de tecnologia atuam de forma significativa na coleta de dados dos usuários das redes, nem sempre com um consentimento plenamente informado, em um contexto de dependência tecnológica característico da cultura digital contemporânea. Nesse cenário, caso um sujeito deseje participar da internet e utilizar determinadas plataformas, torna-se, em alguma medida, vulnerável às condições impostas por essas empresas (Moraes e Testa, 2020).

Essas condições materializam-se, em geral, por meio de termos de uso ou de serviço, que em suas cláusulas envolvem a mineração de dados, permitindo acesso a informações pessoais e de navegação. A mineração de dados se dá em um contexto com uma imensa quantidade de dados circulantes, o que pode ocasionar falhas de processamento nas correlações de sentido.

4.7 Falibilidade

No âmbito da falibilidade algorítmica, destacam-se como conceitos correlatos a imprevisibilidade e erros, o processamento limitado e a volatilidade. Caribé (2017) aborda esse fenômeno ao discutir os sistemas de classificação automática, também conhecidos como auto-tagging, que categorizam e rotulam dados automaticamente. Esses sistemas cometem erros em suas classificações, como em sistemas de avaliação de profissionais que podem resultar em demissões devido a falhas de processamento, por exemplo.

Gillespie (2018b) destaca que a atuação da inteligência artificial, em seus processos decisórios e de resolução de problemas, podem ter erros e fazer associações que nem sempre são corretas. Pant (2022), ao estudar o posicionamento dos desenvolvedores de IA em relação à ética, ressalta a imprevisibilidade desses sistemas, uma vez que algoritmos de aprendizagem de máquina operam com reduzida intervenção humana direta, percorrendo trajetórias próprias e produzindo resultados que, muitas vezes, podem ser questionáveis.

Considerando que, conforme aborda Gillespie (2018), os algoritmos estão sendo constantemente ajustados e modificados, observa-se uma dinâmica de volatilidade, que pode fazer com que, inclusive, os próprios programadores e plataformas percam o controle do desenvolvimento e aprendizagem dessas ferramentas. Essa autonomia gera imprevisibilidade de ação e incerteza (Mittelstadt, 2016). Para Pant (2022), muitas decisões algorítmicas podem violar os direitos pessoais, principalmente quando abordadas a partir da ótica da automatização, a qual está sujeita a falhas e acentuar problemas sociais.

4.8 Viés

No contexto das discussões sobre viés algorítmico, evidenciam-se como conceitos correlatos a ausência de neutralidade e a imparcialidade, que

tensionam a suposta objetividade desses sistemas. Para Santana (2022), os algoritmos não são recursos neutros, uma vez que surgem a partir do momento em que um humano traduz termos e ações para uma linguagem do computador, traduções carregadas de posicionamentos e ideologias. Sendo assim, algoritmos configuram-se como recursos potencialmente enviesados.

Nesse contexto, muitas vezes há limitação de diversidade nos processos de desenvolvimento tecnológico, o que contribui para que visões individuais sejam incorporadas aos algoritmos, de forma intencional ou não. Os indivíduos responsáveis por programar e treinar esses sistemas, bem como por definir os parâmetros de aprendizagem, incorporam valores diversos, tais como racismo, sexismo, concepções de meritocracia e visões estereotipadas, entre outros (Noble, 2018).

4.9 Segregação

No âmbito da segregação algorítmica, destacam-se como conceitos correlatos a discriminação, o preconceito e as desigualdades. Tanto durante a concepção quanto na atuação dos algoritmos nas redes, há a possibilidade de que esses sistemas acentuem práticas discriminatórias e reforcem desigualdades sociais, uma vez que os preconceitos (entendidos como a construção de estereótipos) e as assimetrias estruturais contribuem para a ocorrência de discriminação, ou seja, para o tratamento diferenciado entre usuários, caracterizando processos de segregação.

Para Fernández (2021), um dos principais desafios do ciberespaço é a utilização desses algoritmos que perpetuam e legitimam preconceitos e processos de discriminação. De acordo com Mittelstadt (2016), essa questão viola os direitos de tratamento igual e justo dos indivíduos. Pant (2022) justifica essa questão pela falta de autorreflexão dos próprios desenvolvedores em relação aos seus preconceitos e normas estruturais culturais, que são agregadas aos algoritmos.

A segregação, de fato, influencia os processos comunicacionais, uma vez que pode resultar em um juízo de valor sobre determinadas demandas sociais em detrimento de outras, o que pode invisibilizar temas e agentes importantes e relevantes para um efetivo desenvolvimento social.

4.10 Pós-verdade

No contexto da pós-verdade algorítmica, destacam-se como conceitos correlatos as *fake news*, os *deep fakes*, os golpes, as falsificações e a disseminação de desinformação. Outra discussão que ganha centralidade na contemporaneidade, especialmente após o avanço da inteligência artificial generativa, é aquela que pode ser denominada de pós-verdade algorítmica, a qual engloba o uso de algoritmos para a produção e disseminação de conteúdos enganosos em larga escala.

Conforme Moraes e Testa (2020), as *fake news* consistem em informações inverídicas estrategicamente elaboradas para gerar familiaridade ou aceitação por parte do receptor, criando a falsa impressão de veracidade. Atualmente, os algoritmos são utilizados tanto para a criação quanto para a disseminação dessas informações. Além das notícias falsas, também são produzidos conteúdos manipulados, como imagens, vídeos e áudios sintéticos, o que contribui para a confusão informacional entre os usuários.

García (2021) destaca ainda o uso de softwares baseados em IA para a aplicação de golpes, capazes de imitar vozes e imagens com alto grau de

realismo, dificultando a distinção entre o que é falso e o que é autêntico. Segundo o autor, um estudo conduzido pela Universidade de Oxford e pela Comissão Europeia aponta que 84% dos norte-americanos e 88% dos europeus consideram que a IA deve ser utilizada com cautela, em razão de seu potencial para gerar impactos significativos na sociedade.

5 IMPLICAÇÕES SOCIOTÉCNICAS E ÉTICAS DA GESTÃO ALGORÍTMICA EM REDES SOCIAIS

A análise dos resultados revela que a gestão tecnológica contemporânea nas redes sociais não pode ser dissociada da dimensão ética e social. Os 10 desafios identificados demonstram que os algoritmos deixaram de ser ferramentas meramente técnicas para se tornarem mediadores de processos comunicacionais que influenciam a formação da opinião pública e as subjetividades dos usuários.

Para a gestão, a opacidade e a inescrutabilidade dos algoritmos de “caixa preta” representam um risco estratégico, pois a falta de transparência sobre os critérios de filtragem de conteúdo pode comprometer a autonomia dos indivíduos. Como aponta a literatura, existe uma tensão constante entre os interesses privados das Big Techs (lucro e marketing predatório) e o interesse público.

Nesse cenário, a atuação do gestor tecnológico torna-se central na mediação desses conflitos, exigindo a construção de estratégias que conciliem eficiência econômica e responsabilidade social, evitando que a lógica do capitalismo de vigilância comprometa a confiança dos usuários nas plataformas.

Além disso, a falibilidade e a imprevisibilidade da inteligência artificial exigem que a gestão implemente mecanismos de auditabilidade e intervenção humana. A existência de vieses e práticas de segregação codificadas nos sistemas reforça a necessidade de uma governança algorítmica orientada à diversidade, à equidade e à justiça social, capaz de mitigar os efeitos da desinformação e da pós-verdade e de preservar a integridade dos processos democráticos.

6 CONCLUSÕES

Este estudo cumpriu o objetivo de identificar e sintetizar 10 desafios centrais na gestão de algoritmos em redes sociais on-line, a saber: modulação, arrogância epistêmica, opacidade, interesses privados, capitalismo de vigilância, mineração de dados, falibilidade, viés, segregação e pós-verdade.

A principal contribuição teórica deste trabalho reside na sistematização e categorização crítica desses desafios, realizada a partir da análise de 21 estudos relevantes, o que possibilitou a construção de um panorama estruturado sobre as implicações éticas da atuação algorítmica. Destaca-se, ainda, a elaboração de um quadro sintético (Quadro 2), que organiza e integra os principais conceitos e relações identificadas, contribuindo para a compreensão da complexidade sociotécnica do ambiente digital.

No âmbito prático, os resultados indicam a necessidade de evolução da gestão tecnológica em direção a modelos de governança algorítmica responsáveis. Isso implica a adoção de práticas de transparência ativa quanto ao funcionamento dos sistemas, bem como o desenvolvimento de diretrizes éticas que priorizem a autodeterminação informacional dos

usuários, em contraposição a lógicas estritamente instrumentais de controle e vigilância.

Esta pesquisa apresenta limitações por caracterizar-se como uma revisão bibliográfica de natureza narrativa, o que privilegia a profundidade analítica em detrimento da abrangência estatística. Como agenda futura, sugere-se a realização de estudos empíricos que investiguem os impactos da inteligência artificial generativa na produção e disseminação de conteúdos falsos, como *deep fakes*, bem como a análise da efetividade de regulamentações de proteção de dados sobre as práticas de mineração de dados adotadas pelas plataformas digitais.

REFERÊNCIAS

BARBOSA, M. O. A importância do direito à privacidade digital, redes sociais e extensão universitária. **Fragmentos de Cultura**, Goiânia, v. 24, n. 8, p. 89-97, dez. 2014. Disponível em: <http://dx.doi.org/10.18224/frag.v24i0.3757>. Acesso em: 23 mar. 2022.

BOSTROM, N.; YUDKOWSKY E. A ética da inteligência artificial. **Fundamento - Rev. de Pesquisa em Filosofia**, v. 1, n. 3 - maio - ago. 2011.

CALLEJÓN, F. B. O impacto dos novos mediadores da era digital na liberdade de expressão. **Espaço Jurídico Journal of Law [EJL]**, [S. l.], v. 23, n. 1, p. 179-204, 2022. DOI: 10.18593/ejll.30501. Disponível em: <https://periodicos.unoesc.edu.br/espacojuridico/article/view/30501>. Acesso em: 23 set. 2022.

CARIBÉ, João Carlos Rebello. **Ética na sociabilização mediada por algoritmos**. [S. l.: s. n.], 2017.

CORRÊA, Nicholas Kluge et al. **Worldwide AI Ethics: a review of 200 guidelines and recommendations for AI governance**. 2022. DOI: <https://doi.org/10.48550/arXiv.2206.11922>

DIJCK, Jose Van; POELL, Thomas. Understanding Social Media Logic. **Media and Communication**, v. 1, n. 1, p. 2-14, 2013.

FACEBOOK. **Informações para pessoas que não usam os produtos da meta**. 2023. Disponível em: <https://pt-pt.facebook.com/help/637205020878504>. Acesso em: 18 dez. 2023.

FERNANDÉZ, José Luis. 2021. Hacia el Humanismo Digital desde un denominador común para la Ciber Ética y la Ética de la Inteligencia Artificial. **Disputatio. Philosophical Research Bulletin**, v. 10, n. 17, p. 107-130, 2021.

FERREIRA, D. A. A.; PINHEIRO, M. M. K.; MARQUES, R. M. Termos de uso e políticas de privacidade das redes sociais on-line. **Informação & Informação**, v. 26, n. 4, p. 550-574, 2021. DOI: 10.5433/1981-8920.2021v26n4p550 Acesso em: 23 set. 2022.

GARCÍA, Carlos Lázaro Nodals. Sobre la necesidad de unificación de las iniciativas para un uso ético de la Inteligencia Artificial. **Socialium**, 5(2), 318-334. 2020. <https://doi.org/10.26490/uncp.sl.2021.5.2.880> Acesso em: 23 set. 2022.

GILLESPIE, Tarleton. A relevância dos algoritmos. **Parágrafo**. São Paulo, Brasil, v. 6, n. 1, p. 95-121, jan./abr. 2018

GOMES, Pedro Gilberto. **A Filosofia e a ética da comunicação na midiatização da sociedade**. São Leopoldo. Unisinos, 2006.

HABERMAS, Jürgen. Comunicação política na sociedade mediática: o impacto da teoria normativa na pesquisa empírica. **Líbero**, ano XI, n.21, jun. 2008.

HABERMAS, Jürgen. Political communication in media society: does democracy still enjoy an epistemic dimension? The impact of normative theory on empirical research. **Communication Theory**, v.16, pp. 411-426, 2006.

HABERMAS, Jürgen. **Mudança estrutural da esfera pública: investigações quanto a uma categoria da sociedade burguesa**. Rio de Janeiro: Tempo Brasileiro, 2003.

KARLSON, Michael; VAN COUVERING, Elizabeth; LINDELL, Johan. Publishing, sharing, and spreading online news A case study of gatekeeping logics in the platform era. **Nordicom Review**, v. 43, n. 2, p. 190-213, 2022. DOI: <https://doi.org/10.2478/nor-2022-0012>

LESSIG, Lawrence. **Code 2.0: Code and other laws of cyberspace**. New York: Basic Books, 2006.

MARTINO, Luís Mauro Sá; MARQUES, Ângela Cristina Salgueiro. Ética e Teorias da Comunicação: poder, interações e cultura participativa. **Comunicação e Sociedade**, v. 25, p. 138-153, 2014.

MATTELART, Armand. Comunicação e interesse público. **Comunicação & Educação**. São Paulo, n.16, set/dez 1999.

McQUAIL, Denis. **Atuação da mídia: comunicação de massa e interesse público**. Porto Alegre: Penso, 2012.

MITTELSTADT, B. D.; ALLO, P. A.; TADDEO, M. R.; WACHTER, S.; FLORIDI, L. The ethics of algorithms: Mapping the debate. **Big Data & Society**, 2016.

MITTELSTADT, Brent Daniel. **Principles alone cannot guarantee ethical AI**. 2019. Disponível em: <https://arxiv.org/ftp/arxiv/papers/1906/1906.06668.pdf>

MOORE, George Edward. **Princípios éticos**. São Paulo: Abril Cultural, 1975.

MORAES, João Antonio De, TESTA, Rafael Rodrigues. A Sociedade Contemporânea à Luz Da ética Informacional. **Acta Scientiarum: Human and Social Sciences**, v. 42, n. 3, 2020.

NOBLE, Safiya Umoja. Algorithms of oppression. New York: New York University Press, 2018.

PANT, Aastha et al. **Ethics in ai through the developer's prism: A Socio-Technical Grounded Theory Literature Review and Guidelines.** arXiv preprint, 2022.

PEDROSA, Clara Bonaparte; JÚNIOR, José Alfredo De Oliveira Baracho. Bolha Informacional E Mídias Sociais: Desafios para as eleições na era da sociedade da informação. **Revista Thesis Juris**, v. 10, n. 1, p. 148-164, 2021.

RAMÍREZ, D. G., JIMÉNEZ, D. V. Los Impactos De La Ideología Técnica Y La Cultura Algorítmica En La Sociedad: Una Aproximación Crítica. **Revista De Estudios Sociales**, n. 71, p. 15-27, 2020.

SANTAELLA, Lucia. **A pós-verdade é verdadeira ou falsa?** São Paulo: Estação das Letras e Cores, 2018.

SANTAELLA, Lucia; KAUFMAN, Dora. **Os dados estão nos engolindo?** *Civitas*, v. 21, n. 2, p. 214-223, 2021. DOI: <https://doi.org/10.15448/1984-7289.2021.2.39640>

SANTANA, Ramon Davi; NEVES, Barbara Coelho. Entre filtros e bolhas: a modulação algorítmica na sociedade pós-panóptica. **Logeion**, v. 8, n. 2, p. 47-64, 2022. DOI: <https://doi.org/10.21728/logeion.2022v8n2.p47-64>

SILVEIRA, Sergio Amadeu da. A noção de modulação e os sistemas Algorítmicos. **Dossiê**, v. 3, n. 6, 2019. DOI: <http://dx.doi.org/10.31657>

STROPPIA, Tatiana. **Plataformas Digitais e Moderação de Conteúdos:** por uma regulação democrática. Belo Horizonte: Fórum Ltda, 2021.

FOUCAULT, Michel. **Vigiar e punir:** nascimento da prisão. 14. ed. Petrópolis: Vozes, 1996.

VIGIL, José L. García. **Reflections around ethics, human intelligence and artificial intelligence**, 2021

VRIES, K. Identity, profiling algorithms and a world of ambient intelligence. **Ethics and Information Technology**, v. 12, n. 1, p. 71-85, 2010.

WE ARE SOCIAL. Digital 2022: another year of bumper growth. 2022. Disponível em: <https://wearesocial.com>. Acesso em: 21 set. 2019.

ZUBOFF, Shoshana. **A era do capitalismo de vigilância:** a luta por um futuro humano na nova fronteira do poder. Rio de Janeiro: Intrínseca, 2019.